ABSTRACT

Title of dissertation: A BIOPHYSICAL STUDY OF CLATHRIN UTILIZING LIGHT SCATTERING, NEUTRON SCATTERING AND STRUCTURE BASED COMPUTER MODELING

Matthew Lee Ferguson, Doctor of Philosophy, 2007

Dissertation directed by: Dr. Ralph Nossal
National Institute of Child Health and Human Development

and

Professor Wolfgang Losert
Department of Physics

A principal component in the protein coats of certain post-golgi and endocytic vesicles is clathrin, which appears as a three-legged heteropolymer (known as a triskelion) that assembles into polyhedral baskets principally made up of pentagonal and hexagonal faces. In vitro, this assembly depends on the pH, with baskets forming more readily at low pH and less readily at high pH. We have developed procedures, based on static and dynamic light scattering, to determine the radius of gyration, R_g, and hydrodynamic radius, R_H, of isolated triskelia under conditions where basket assembly occurs. Calculations based on rigid molecular bead models of a triskelion show that the measured values can be accounted for by bending of the legs and a puckering at the vertex. We also show that the values of R_g and R_H measured for clathrin triskelia in solution are qualitatively consistent with the conformation of an individual triskelion that is part of a "D6 barrel" basket assembly measured by cryo-EM tomography.

We extended this study by performing small angle neutron scattering (SANS) experiments on isolated triskelia in solution under conditions where baskets do not assemble. SANS experiments were consistent with previous static light scattering experiments but showed a shoulder in the scattering function at intermediate q-values just beyond the central diffraction peak (the Guinier regime). Theoretical calculations based on rigid bead models of a triskelion showed well-defined features in this region different from the experiment. A flexible bead-spring model of a triskelion and Brownian dynamics simulations were used to generate a time averaged scattering function. This model adequately described the experimental data for flexibilities close to previous estimates from the analysis of electron micrographs.

A BIOPHYSICAL STUDY OF CLATHRIN UTILIZING LIGHT SCATTERING, NEUTRON SCATTERING AND STRUCTURE BASED COMPUTER MODELING

by

Matthew Lee Ferguson

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2007

Advisory Committee:
Dr. Ralph J. Nossal, Co-Chair/Advisor
Professor Wolfgang Losert, Co-Chair/Advisor
Dr. Dan L. Sackett
Professor Sandra Greer
Professor J. Robert Dorfman
Professor Robert W. Gammon

DEDICATION

I would like to dedicate this dissertation to the memory of
Douglas Wright Ferguson

ACKNOWLEDGMENTS

First, I would like to thank Ralph Nossal and Wolfgang Losert. Ralph Nossal provided me the problem addressed in this dissertation and the opportunity and support to work on it. He was always encouraging and supportive throughout while allowing me a great amount of independence and creativity. I thank Wolfgang Losert for his support, mentoring, and continued advice and encouragement over the course of my education. I would also like to thank both Ralph and Wolfgang for the patience they both have shown me.

This work is the culmination of the efforts of a large number of other people who contributed greatly. I would especially like to thank Dan Sackett and Hacene Boukari, both of whom patiently helped me with their time and knowledge and experience on this project. Kondury Prasad and Eileen Lafer made this project possible by purifying the protein, giving generously whenever we asked for new material, over the course of the last three years. Without their help, hard work, and sacrifice none of the work I present here would have been possible. Rudolfo Ghirlando provided the generous gift of his light scattering instrument and his time and experience in helping me to align and use the instrument. Albert Jin made early light scattering and neutron scattering measurements on clathrin and was very kind and helpful in getting me started on this project. Chris Stanley generously performed our first SANS experiment with our clathrin sample on his own beamtime. Susan Krueger assisted me in writing the SANS beamtime proposal, helping to take the measurements, along with Ralph and Hacene. She also helped me to do the modeling of the SANS data by smearing my models and suggesting the reason for the disagreement with rigid model calculations. Peter Schuck helped me by performing analytical ultracentrifugation measurements. Richard Pastor provided me with reprints of his papers on computer simulations, provided advice on writing and checking the code, and performed a hydrodynamics calculation for me to check the translational-rotational coupling of my triskelion model using his own hydrodynamics program. Jack Douglas has provided me with much support and encouragement throughout, but was especially helpful to me with advice and instruction on using ZENO, as well as showing me where to find the code and documentation to run it myself. Marc Mansfield performed a great number of preliminary calculations for me using ZENO before I could run the code, Jose Garcia de la Torre helped me by providing details about HYDRO through direct and email correspondence, Ferenc Horkay provided me with advice on analysis of SANS data on clathrin as well as light scattering experiments on polymers, which I do not discuss here.

I would also like to thank Don Martin, Dan Lathrop, Jeff Urbach, Sandra Greer, Jermey Matthews, Jeff Forbes, Bob Dorfman, Ed Ott, Daniel Wojcik, Vincent Ferriero, and Daniel Lanterman, all of whom helped me at one time or another in previous research projects which are not discussed here.

Thanks to Ron Skupsky, Ariel Michelman-Ribiero, Jennifer Galanis, Daniel Harries, Brian Todd, Horia Petrache, Jason Derouchey, Adrian Parsegian and Bob Bonner, who taught me many things in daily conversation. For supporting the graduate student community at NIH, I would like to thank Richard McGee and Mary Delong of the NIH Graduate partnership program and Brenda Hanning of NICHD.

I would like to thank the people in the office for all of their help. They are Raci Say, Jennifer Cooper, Visal Peang Sam and Erlinda Inejosa-Ortanez.

From the University of Maryland, I thank Kyuyong Lee, Gene Kwon, Sebastian Bailard, Masahiro Toya, Sander Van der Meer, Mike Knewey, Erin Rericha, Jay Pyle, Nancy Boone, Carol Belamy, Jane Hessing, Nicholas Chant, Linda Ohara, Berne Koslowski, Lubna Rana and Wendel Hill

Also for friendship, moral support and personal advice, Breno Imbiriba, David Fiske, Violeta Prieto, Dan Sisan, Heather Lindsay, Tim Stasevich, Ayumi Yamada, Carlos Sanchez, Brad and Shirley Winter, Woody Shew, Mel Eck, Jo Ann Thompson, Greg Smith, Anthony Saldino, Ferenc Szalma, Timothy Spears, Bruce Miller, Jon Teague, Madelaine Monroe, Jonas Rosenthal, Nathan and Jamie Morris, Diana Irvin, Shanti, Daniel, and Adelaide Avazpour, Matt Stringer and Cristi Lemon.

And my mother and father, Janna K. and Gary W. Ferguson.

During the course of this work, I was supported by an intramural fellowship (a pre-doctoral IRTA) from NICHD.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

1.1 Foreword

Developing a better understanding of how cells operate requires advances in the understanding of how basic cellular functions such as mitosis (cell division), exocytosis (molecular secretion), and endocytosis (molecular internalization) are performed at the molecular level[1]. Each of these processes is driven by the collective function of many different cellular components having the form of cytosolic proteins, membrane proteins, lipids and nucleic acids.

In different types of cells, many of these components have been identified. Their roles in cellular functions can be discerned in many ways. One common way is by "removing" a component from a cell to see how (or whether) a process is affected. If a process is affected, then the missing component may be involved in that process in some way. This can be done by producing a knockout mouse, making a mutation in the genome, or more recently by RNA interference (RNAi). The static structure of some components may be determined by X-ray crystallography or Nuclear Magnetic Resonance Spectroscopy (NMR). The molecule must be crystallized for X-ray and

have a low mass (< 40 kDa[1]) for solution NMR[2]. Cryo-EM tomography can be used to reconstruct some larger macromolecular structures and assemblies. By identifying and visualizing the basic building blocks of some parts of the cellular machinery, we hope to get some idea of how each component performs its assigned function within the cell.

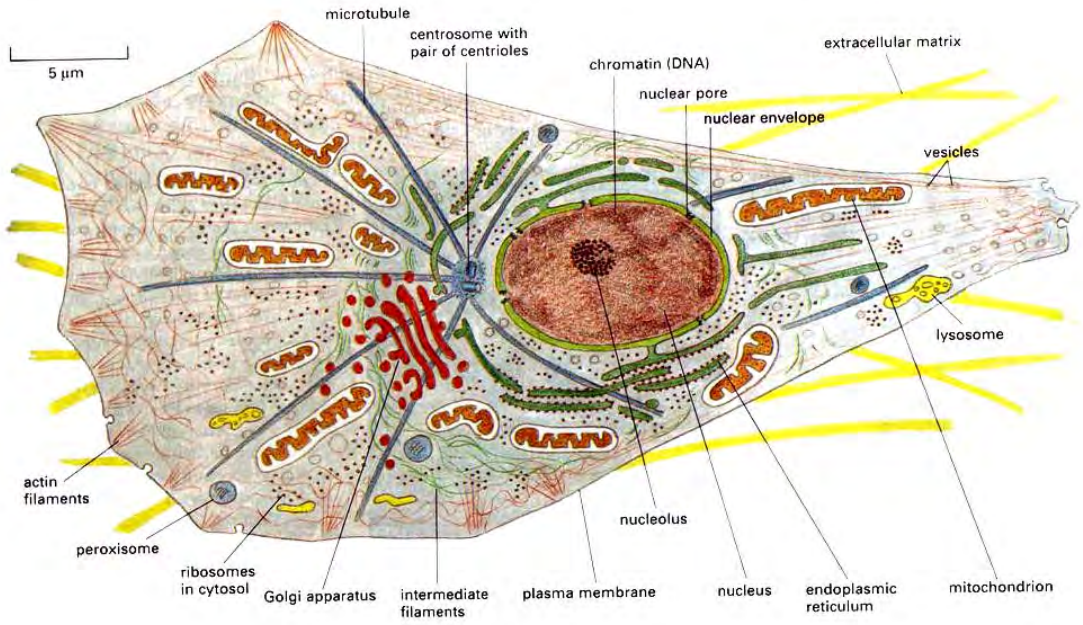

Fig. 1.1: "The major features of eucaryotic cells. The drawing depicts a typical animal cell, but almost all the same components are found in plants and fungi and in single-celled eucaryotes such as yeasts and protozoa." Figure and caption from *Molecular Biology of the Cell*[4, Fig 1-31, p. 29]

[1] 1 kDa $= 10^3$ Dalton $= 10^3$ atomic mass units $= 1/N_A$ kilograms (where N_A is Avagadro's Number)

[2] Recently much larger structures ($\sim$900 kDa) have been obtained using NMR but these structures require a high degree of symmetry. See Ref. [2] and [3].

In eucaryotic[3] cells, membranes composed of lipid bilayers and integral membrane proteins make up the outer plasma membrane and subdivide the cell into different compartments or organelles. (See Fig. 1.1.) The endocytic pathway allows cells to ingest macromolecules from outside of the cell. The secretory pathway allows eucaryotic cells to regulate the delivery of molecules (such as newly synthesized proteins) from the cell interior to the cell exterior. These processes are collectively called intracellular "trafficking".

The clathrin machinery (which is the system under study) is involved in both the endocytic and secretory pathways. The protein, clathrin, is a heteropolymer composed of a 192 kDa heavy chain and a variable, ca. 23-27 kDa, associated light chain [5]. It occurs in nearly all eucaryotic cells. It is a component of the machinery that the cell uses to move specific molecules across lipid bilayer membranes, in the process generating a clathrin coated vesicle which is composed of a molecular lattice surrounding an inner membrane shell.

Cells also use clathrin independent pathways for trafficking, including the use of the proteins called coat protein complex I and II (COPI and COPII) in the secretory pathway, and the protein caveolin in the endocytic pathway[1, ch. 14 p. 472]. These other types of intracellular trafficking are known to have specific roles within the cell,

[3] Eucaryotic cells are a large class of cells whose defining characteristics are that genetic information is contained in a nucleus, and the cells are divided up internally into different compartments by lipid bilayer membranes. Animal kingdoms belonging to the domain Eucaryota include Animalia, Fungi, Plantae, and Protista. Prokaryotic cells are evolutionarily older and are not divided up in this way. Prokaryota includes the domains of Bacteria and Archaea. In prokayotes genetic information is spread throughout the cell.

as does clathrin. It is not currently understood why the cell requires more than one pathway for trafficking or why they are different[6]. Fig. 1.2 shows the utilization of these coat proteins in different aspects of the intracellular trafficking pathway.

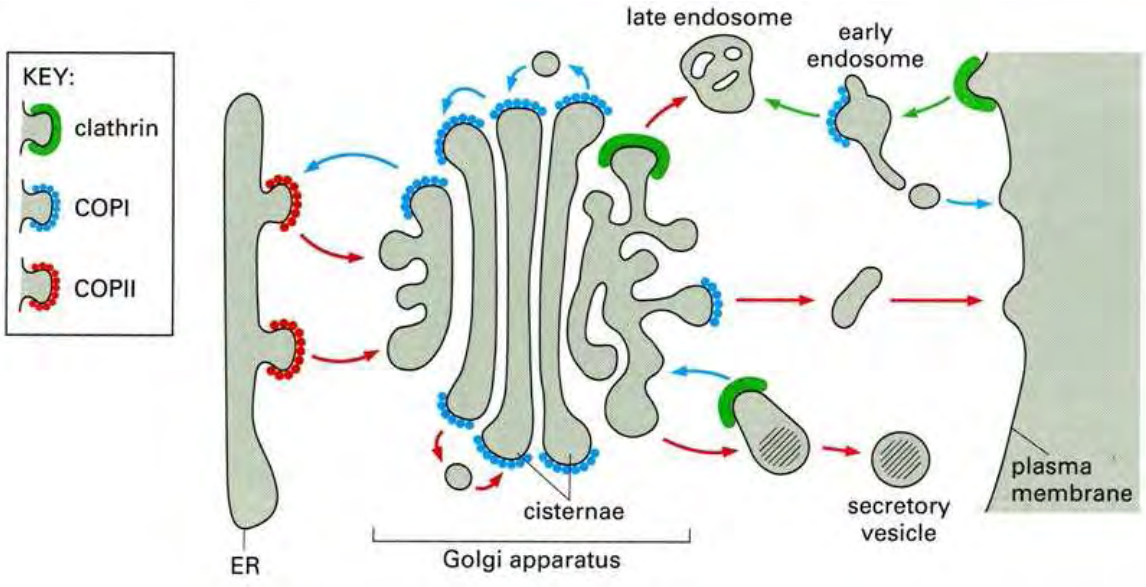

Fig. 1.2: "Utilization of different protein coats in intracellular trafficking. Different coat proteins select different cargo and shape the transport vesicles that mediate the various steps in the secretory and endocytic pathways. Many differentiated cells have additional pathways beside those shown in this figure, including for example, a specialized recycling pathway for proteins of synaptic vesicles in the synapses of neurons." Figure and Caption from *Molecular Biology of the Cell*[4, Fig 13-5 p. 716]

All of the proteins listed above, clathrin, COPI, COPII and caveolin operate by coating small (∼60nm) vesicles that carry cargo through the cell. Some specific and important examples of clathrin function are LDL intake, insulin degradation and

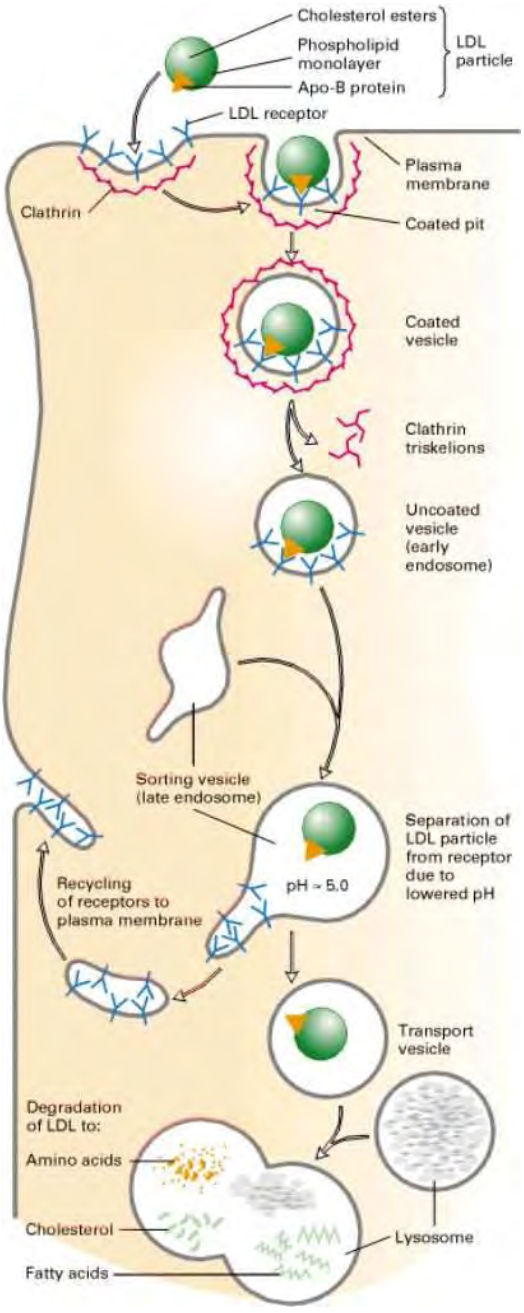

Fig. 1.3: "Fate of an LDL particle and its receptor after endocytosis. After an LDL particle binds to an LDL receptor on the plasma membrane, the receptor-ligand complex is internalized in a clathrin-coated pit that pinches off to become a coated vesicle. The clathrin coat then depolymerizes to triskelia, resulting in an early endosome. This endosome fuses with a sorting vesicle, known as a late endosome, where the low pH (5) causes the LDL particles to dissociate from the LDL receptors. A receptor-rich region buds off to form a separate vesicle that recycles the LDL receptors back to the plasma membrane. A vesicle containing an LDL particle may fuse with another late endosome but ultimately fuses with a lysosome to form a larger lysosome. There, the apo-B protein of the LDL particle is degraded to amino acids and the cholesterol esters are hydrolyzed to fatty acids and cholesterol. Abundant imported cholesterol inhibits synthesis by the cell of both cholesterol and LDL receptor protein. The same pathway is followed by other ligands, such as insulin and other protein hormones, that are internalized by receptor-mediated endocytosis and degraded in the lysosome." Figure and Caption from *Molecular Cell Biology*[7]

synaptic vesicle formation in neurons[4]. More generally, clathrin-coated vesicles are involved in the trafficking of proteins between specific organelles within the cell and the recycling/recovery of constituents of the outer cell membrane. A schematic of coated vesicle formation is shown in Fig. 1.3.

Three clathrin molecules join at a common hub to form a three-legged "triskelion", which is the basic building block of the lattice. Each leg is approximately 30 Å thick[8] and 520 Å [9, 10] in length, and ends in a globular "terminal domain" of radius 50 Å[11, 8]. By convention, a leg is divided into a proximal segment adjoining the central hub, a distal segment, and a linker region that connects with the terminal domain. A clathrin heavy chain runs the entire length of the leg, with a light chain attached near the common hub (see Fig. 2.1).

In vitro, the triskelia assemble into polyhedral "baskets" composed of pentagonal and hexagonal faces, mimicking the structures seen on the outside of endocytic vesicles and the inner plasma membrane of a cell[13] (shown in Fig. 1.4). One can dissociate baskets or uncoat vesicles by changing the properties of the solution in which they are suspended (e.g., by changing pH and ionic strength) [14] or by adding an uncoating protein (called Hsp70 ATPase [15]).

The structure assumed by a triskelion in a basket has been measured by cryo-EM tomography[16]. (This structure is shown in Fig. 2.3.) The solution structure of an isolated triskelion cannot easily be imaged directly, but must be inferred from

[4] Synaptic vesicles carry neurotransmitter for rapid release into a synaptic junction by a firing neuron.

Fig. 1.4: Electron micrograph of inner plasma membrane of a cell. A deep etch electron micrograph of the inner surface of the plasma membrane of a cell undergoing endocytosis. Clathrin coated pits and vesicles are clearly visible.[12, from J. Heuser]

indirect, although noninvasive, measurements. Those measurements and the analysis done to interpret them are the subject of this thesis.

Static light scattering (SLS) provides a measure of particle size and dynamic light scattering (DLS) provides a measure of particle mobility (or diffusion). The first aim of our study was to use these two techniques to determine information about triskelion structure in solution. These measurements could then be compared with expected values based on the structure of a triskelion within an assembled clathrin basket, which has been determined by high resolution cryo-EM tomography[16]. This was done to determine whether the puckered shape of the triskelion in the lattice preexists before the assembly of the clathrin lattice, or whether assembly into the lattice produces this triskelial shape.

The second aim of our study involves understanding of the flexibility of the triskelion leg. This is of crucial importance to measurement of the solution structure of the triskelion, the first aim. If the triskelion leg is inflexible, then each individual triskelion must necessarily have a predetermined shape allowing it to self assemble into a basket under appropriate conditions. If the triskelion leg is very flexible then a unique triskelial structure does not exist before the assembly into the basket. In this case the puckered structure for a clathrin triskelion would exist only because of geometry imposed by the lattice of a clathrin basket. The experimental technique that is used to investigate triskelial flexibility is small angle neutron scattering (SANS).

SANS was used in this study in order to extend static light scattering measurements using a probe of smaller length scales (neutrons used had a wavelength of 5.5 Å vs. a wavelength of 5145Å for our SLS). It was not known *a priori* what we

would find by SANS, but it was assumed that these measurements would support or contradict our rigid bead models of clathrin which were used to fit light scattering data of the previous study.

1.2 Outline of Thesis

Chapter 2 includes background information on the biological molecule under study as well as a review of the basic physics underlying the experimental and computational techniques that were used to interpret the results of my experiments.

Chapter 3 describes light scattering experiments that were performed on isolated clathrin triskelia in solution along with analysis that was done to interpret the results. Both Dynamic Light Scattering (DLS) and Static Light Scattering (SLS) techniques were used to measure scattered laser light from purified protein in solution. One of the questions asked was: Does the clathrin triskelion have the same shape when it is isolated in solution as it has when assembled into a clathrin basket as measured by cryo-EM tomography? (The cryo-EM structure of a clathrin basket as measured by Fotin et al.[16] is shown in Fig. 2.3.) The results of my experiments were compared to different models of clathrin that were based on previous structural studies.

Chapter 4 describes small angle neutron scattering (SANS) experiments that were performed on isolated clathrin triskelia in solution, as well as analysis using rigid models. Small angle neutron scattering experiments measured neutrons scattered from purified clathrin triskelia in solution. One of the questions asked was: Can we

gain additional information about clathrin's structure in solution by doing SANS that is not available through SLS and DLS? The results of the experiments were compared to calculations using the Debye equation and different models of clathrin based on previous structural studies (including those of Chapter 3).

Chapter 5 describes Brownian dynamics simulations which take into account the flexibility of a clathrin triskelion molecule. The simulation is used to calculate a time-averaged scattering function. The calculated scattering function is then compared with SANS data from experiments described in Chapter 4. We find that this analysis provides an estimate of the persistence length of the clathrin legs.

Chapter 6 summarizes our study and discusses future experiments.

2. BACKGROUND

2.1 Clathrin Structure and Function

Clathrin is an intriguing candidate for detailed molecular investigation. The striking electron micrographs taken by John Heuser and Jim Harrison in 1983[12, 17] (similar to that seen in Fig. 1.4) showed the polygonal coat of the inner membrane of a cell where clathrin mediated endocytosis occurs. When clathrin triskelia were isolated from coated vesicles, electron micrographs showed the three armed pinwheel like structure of the triskelion shown in Fig. 2.1a.

Images and structures of the components of this complex macromolecule have improved, but a simple understanding of how it works is elusive. For example, it is not known whether clathrin causes the membrane to bend as was originally thought, or whether it recruits proteins to the region that cause the membrane to bend as some new evidence has shown[18, 19].

If the legs of a triskelion were to be characterized as a polymer, a good description might be that they were semi-flexible, i.e. bending on length scales comparable to the length of a single leg. They have both definite structural attributes and flexibility[9, 10]. This may or may not be a necessary property for the function they

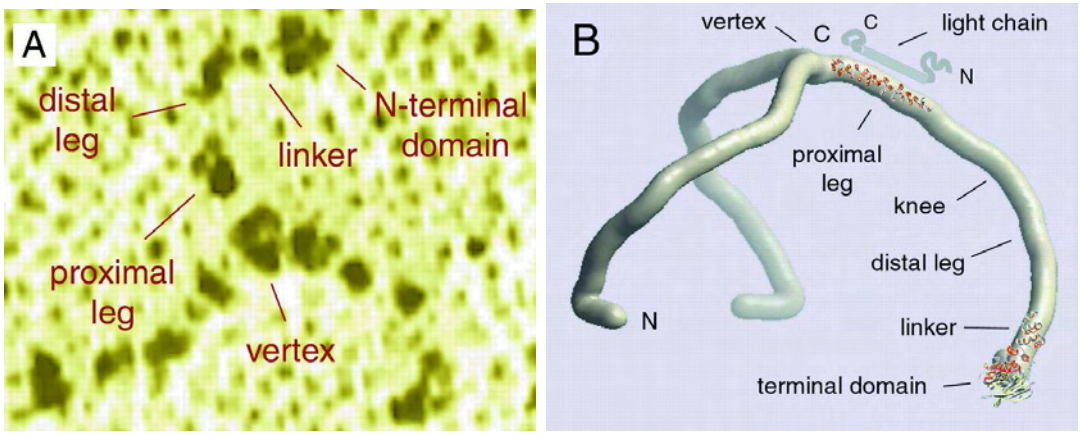

Fig. 2.1: a) "An electron micrograph showing a clathrin triskelion."[5, Fig. 4a]

b) "A diagram of the lattice structure of a clathrin triskelion, measured by cryo-EM tomography of a D6 basket assembled from Clathrin and AP2." Figure and Caption from Kirchhausen et al.[5, Fig. 4b]

perform, but the flexible nature of triskelia requires special approaches in order to study their structure.

2.1.1 Clathrin is an Unstable Protein

A common way to produce a protein, if its amino acid sequence is known, is to express it in bacteria or yeast[1] cells. Clathrin cannot be produced in this fashion because it will not fold properly without the presence of eucaryotic chaperone proteins[20]. Also, for unknown reasons, purified clathrin triskelia "degrades" after approximately two weeks. Storage at low temperature ($-80°$ C) does not change this process of degradation[21]. Therefore, fresh protein must be acquired and purified

[1] These are two cell types that are easily produced and are viable and reproductive after this kind of genetic modification.

from bovine brain matter on a regular basis. Experiments generally must be completed within two weeks of purification, or the protein will lose its ability to function.

2.1.2 Previous Structural Measurements of Clathrin Triskelia

The subunits of supramolecular clathrin structures are descriptively called triskelia, since each clathrin triskelion is made up of three legs. Early images of clathrin triskelia came from rotary shadowed electron micrographs of triskelia adsorbed onto cleaved mica surfaces.[17] (See Fig. 2.1a)

Each "leg" of a clathrin triskelion is made up of two polypeptide chains. The largest polypeptide chain, called the clathrin heavy chain or CHC, is 1675 amino acids long and has a molecular weight of 192 KDa.[5, p. 707] The C-terminus of the CHC is a trimerization domain, where the three legs join. A large part of the CHC has a coiled coil zig-zag structure of α-helices[22], somewhat reminiscent of a spring. The N-terminus is a seven blade β-propeller[11]. This is a binding domain for associated proteins also involved in clathrin-mediated endocytosis. Pieces of both of these parts of the CHC have been crystallized and the structures determined at high resolution by X-ray diffraction (see Fig. 2.2).

Attached to the CHC near the trimerization domain is the clathrin light chain (or CLC). This polypeptide has two forms, A and B. CLC A contains 243 amino acids and has a molecular weight of 26.7 KDa. CLC B has 210 amino acids and a molecular weight of 23.0 KDa[2]. These different types of CLC can attach onto a single

[2] Molecular weights were calculated from the known protein sequences by the online tool and database: Expasy ProtParamTool. (http://ca.expasy.org/tools/protparam.html) The protein sequences

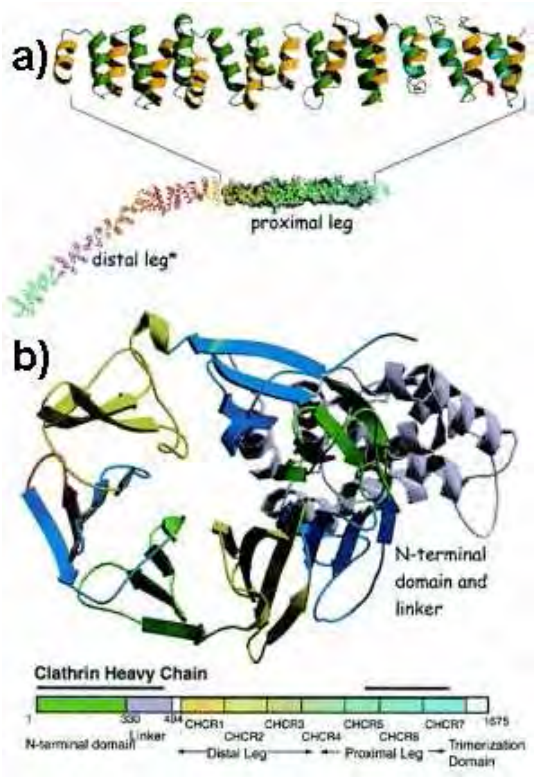

Fig. 2.2: Partial structures of the clathrin heavy chain. (a) Clathrin heavy chain proximal leg (1210-1516) and (b) N-terminal domain and linker. From Wakeham et al.[8]

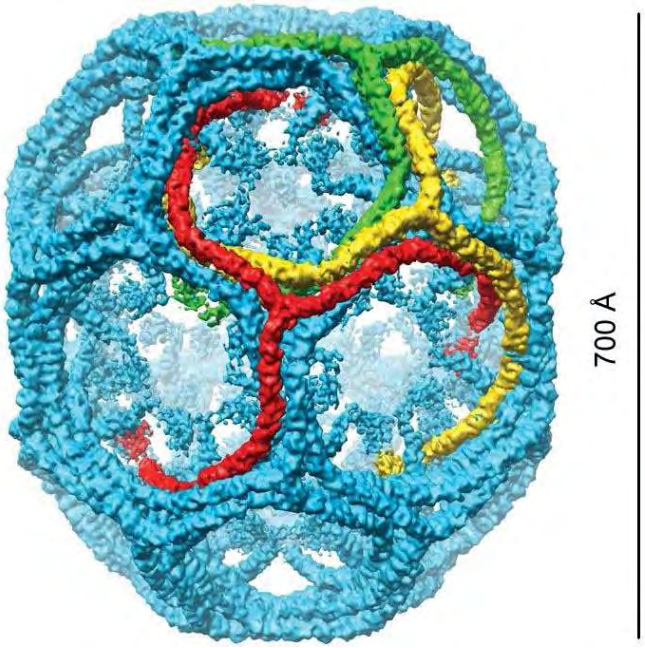

Fig. 2.3: "A high resolution cryo-EM reconstruction of a clathrin basket. Image reconstruction of a clathrin D6 basket (only the heavy chains are shown) at 7.9Å resolution." Figure and Caption from Fotin et al.[16, Fig. 2]

triskelion in any combination AAA, AAB, ABB, or BBB. They share about 60% sequence identity and are thought to regulate triskelion basket assembly and affect CHC trimerization[23]. CLC A and CLC B are both believed to be unstructured when not attached to the CHC.

The entire clathrin triskelion is a hetero hexamer of CHC, CLC A and CLC B. It average molecular weight is ∼650 KDa. Each arm stretches out from its vertex in what has been described as a "spider" or "pinwheel" like structure. The average length of the arm is estimated to be ∼520 Å long. The average width is ∼30 Å[9].

used were: CHC: NP_776448 CLC A: NP_776447 CLC B: NP_776702

from *Bos Taurus*

2.1.3 Clathrin Coated Vesicles and Clathrin Baskets

Vesicles isolated from different types of cells have characteristic sizes and associated clathrin coats that can be related to specific polyhedra, made up of pentagonal and hexagonal faces[13]. All polyhedral structures have a single triskelion at each vertex.

When isolated in solution, under special conditions clathrin triskelia will self-assemble to form empty, closed polyhedral structures called "cages" or "baskets". These have a morphology similar to that of the clathrin lattice surrounding coated vesicles. The structures observed include tetrahedral and cubic[24] as well as dodecahedral and icosahedral polyhedra. In the presence of an adaptor protein, AP2, a structure called a D6 basket has been observed to occur. Recently, this assembly has been studied by cryo-EM tomography at a resolution great enough to distinguish individual clathrin triskelia within the structure[16]. (This structure is shown in Fig. 2.3) This measurement suggests that, in a lattice, triskelia do not assume a conformation that is flat and planar like the structure shown in electron micrographs, but rather the structure has an extreme umbrella-like curvature exceeding the curvature of the basket structure it forms. The shape of an individual triskelion when incorporated into such a lattice structure is shown in Fig. 2.1b[16] and is highlighted in red in Fig. 2.3.

2.2 Scattering from Dilute Macromolecules in Solution

In the following section, I will derive basic equations that will be used in later chapters to extract information from experimental measurements from clathrin in solution. The objective of this section is to review the basic equations governing the scattering of light and neutrons from macromolecules in solution. Interpretations of experimental results that will be presented later will be dependent on this theory.

2.2.1 Outline of Section

The first result derived, the Debye equation, describes the diffraction pattern of waves scattered from a molecule in solution, when the target is large compared to the wavelength of the incident radiation. In this derivation, the molecule is approximated by an arrangement of identical point scatterers. The resulting diffraction pattern is due to constructive and destructive interference of waves scattered from different points within the molecule. The second result, the Guinier Approximation, is derived from the Debye equation and is a simplification of the equation in the small-angle limit or the limit of large but finite wavelength. The third result applies to a molecule undergoing diffusion and relates the motion of the molecule to fluctuations in the scattered light intensity and the temporal intensity autocorrelation function.

2.2.2 Origin of Scattered Intensity of Light from Macromolecules in Solution

Incident light sets up a vibration between the positively charged and negatively charged elements of a molecule or dielectric particle, creating an oscillating dipole.

Such a dipole is itself a source of electromagnetic radiation. The secondary radiation has the same frequency as the incident radiation but is 180° degrees out of phase with it (note the minus sign in Eq. 2.1). So the electric dipole moment, $\vec{p}$, induced at a position $\vec{r}_0$ by a traveling electromagnetic wave, $\vec{E_0}$, incident on a point particle of polarizability α would be given by

$$\vec{p} = -4\pi\epsilon_0\alpha\vec{E}_0 e^{i\vec{k}_0\cdot\vec{r}_0} \tag{2.1}$$

where the polarizability, α, has units of volume and ϵ_0 is the electric permittivity of free space. The electric field of the dipole radiation, $\vec{E}_S$, radiating along a direction $\hat{n}$ from such a dipole is given at a position $\vec{r}$ (far from the particle so $\vec{r} \gg \vec{r}_0$) by[25]

$$\vec{E}_S = \frac{k_0^2}{4\pi\epsilon_0}\frac{e^{ik_0\hat{n}\cdot(\vec{r}-\vec{r}_0)}}{|\vec{r}-\vec{r}_0|}(\hat{n}\times\vec{p})\times\hat{n}$$

using the approximation that $k_0\hat{n}\cdot(\vec{r}-\vec{r}_0) \approx k_0 r - k_0\hat{n}\cdot\vec{r}_0$, Eq. 2.1, and by defining the scattered wave vector as $\vec{k}_S = k_0\hat{n}$ and the scattering vector as $\vec{q} = \vec{k}_0 - \vec{k}_S = \vec{k}_0 - k_0\hat{n}$. This expression can be approximated as

$$\begin{aligned}\vec{E}_S &\approx -k_0^2\alpha e^{i\vec{k}_0\cdot\vec{r}_0}\frac{e^{ik_0 r}}{r}e^{ik_0\hat{n}\cdot\vec{r}_0}(\hat{n}\times\vec{E}_0)\times\hat{n} \\ &\approx -k_0^2\alpha\frac{e^{ik_0 r}}{r}e^{i\vec{q}\cdot\vec{r}_0}(\hat{n}\times\vec{E}_0)\times\hat{n}. \end{aligned}\tag{2.2}$$

Moreover, $|(\hat{n}\times\vec{E}_0)\times\hat{n}|$ becomes $|\vec{E}_0|\sin(\theta_1)$, where θ_1 is the angle between the incident wave's polarization (also the direction of the induced electric dipole) and the scattering plane. The scattering plane is defined by the incident wave vector, $\vec{k}_0$, and outgoing wave vector, $\vec{k}_S = k_0\hat{n}$. Since in our light scattering experiments the incident laser light is vertically polarized, $\theta_1 = 90°$. This assumes that the molecule under

study is an isotropic scatterer (the polarization, $\vec{p}$, is aligned with and proportional to incident electric field, $\vec{E}_0$ as in Eq. 2.1). The scattered intensity is thus given by

$$I_S = |\vec{E}_S^* \cdot \vec{E}_S| = \frac{\alpha^2 k_0^4 \sin^2(\theta_1)}{r^2} I_0. \tag{2.3}$$

Eq. 2.3 was first derived by Lord Rayleigh in 1871 and describes light scattered from a simple point dipole or scattering from a molecule whose dimension is small compared to the wavelength of the incident radiation.

Since we do not scatter from particles in a vacuum but from particles in a solvent, we are interested in the difference in polarizability of the particle with the surrounding solvent. The polarizability relative to the medium, α, can be related to the index of refraction by the equation $4\pi\rho\ \alpha = n^2 - n_0^2$ where n is the index of refraction of the particle, n_0 is the index of refraction of the surrounding medium and ρ is the number of scattering particles per unit volume[26].

2.2.3 Origin of Scattering of Neutrons from Macromolecules in Solution

Neutrons interact in a different manner than electromagnetic radiation. Instead of inducing an electromagnetic dipole, neutrons interact with the spins of an atomic nucleus through the short ranged strong nuclear force. Since the range of this force is so short relative to the wavelength of neutrons we are working with (10^{-15} m vs. 10^{-10} m) the neutron-nucleus interaction can be approximated by a delta function (Fermi's pseudo-potential) which, for a single nucleus j centered at position $\vec{r}_j$ with scattering length b_j, is given by

$$V_j(\vec{r}) = \left(\frac{2\pi\hbar^2}{m}\right) b_j \delta^3(\vec{r} - \vec{r}_j),$$

where $\delta^3(\vec{r}-\vec{r}_j)$ is a Dirac delta function centered at $\vec{r}_j$. Thus, for a collection of N nucleons, the scattering potential would be

$$V(\vec{r}) = \sum_{j=1}^{N} V_j(\vec{r}) = \left(\frac{2\pi\hbar^2}{m}\right)\sum_{j=1}^{N} b_j\delta^3(\vec{r}-\vec{r}_j).$$

The scattering length, b_j, must be measured for each element and can be found on the National Center for Neutron Research webpage[3] for "thermal" neutrons (having energies close to room temperature, corresponding to a wavelength greater than 1Å). It was shown by Fermi that this pseudo-potential causes the same scattering as the actual interaction but is weak enough to be used in the perturbation expansion of Born (the Born Approximation)[27, 28].

Far from a scatterer positioned at the origin, a scattered plane wave can be written in the form

$$\Psi_S(\vec{r},t) = f(\Omega)\frac{e^{ik_0 r}}{r}. \tag{2.4}$$

For a weak potential $V(\vec{r})$ in the Born approximation, the scattering wave amplitude, $f(\Omega)$, is given by[29]

$$\begin{aligned} f(\Omega) &= -\frac{m}{2\pi\hbar^2}\int e^{i(\vec{k}_0-\vec{k}_S)\cdot\vec{r}}V(\vec{r})d\vec{r} \\ &= -\frac{m}{2\pi\hbar^2}\int e^{i\vec{q}\cdot\vec{r}}V(\vec{r})d\vec{r} \\ &= -\sum_{j=1}^{N} b_j e^{-i\vec{q}\cdot\vec{r}_j}. \end{aligned} \tag{2.5}$$

For a single point dipole in light scattering

$$f(\Omega) = \alpha k_0^2|\vec{E}_0|\sin(\theta_1)e^{i\vec{q}\cdot\vec{r}_0} \tag{2.6}$$

[3] http://www.ncnr.nist.gov/resources/sldcalc.html

so, for many dipoles, a light scattering expression appears, in analogy to Eq. 2.5, as

$$f(\Omega) = \sum_{j=1}^{N} \alpha k_0^2 |\vec{E}_0| \sin(\theta_1) e^{i\vec{q}\cdot\vec{r}_j}. \tag{2.7}$$

The probability of an incident plane wave of wave vector $\vec{k}_0$ being scattered by a weak potential $V(\vec{r})$ to become an outgoing plane wave with outgoing wavevector $\vec{k}_S$ is given by

$$\begin{aligned} \frac{d^2\sigma}{d\Omega} &= |f(\Omega)|^2 \\ &= \sum_{i=1}^{N}\sum_{j=1}^{N} b_i b_j e^{i\vec{q}\cdot(\vec{r}_i - \vec{r}_j)} \\ &= \sum_{i=1}^{N}\sum_{j=1}^{N} b_i b_j e^{i\vec{q}\cdot\vec{r}_{ij}}. \end{aligned} \tag{2.8}$$

Since $V(r)$ is a real valued potential, $b_j^* = b_j$.

Neutrons have a very strong interaction with the protons in hydrogen atoms. If water is used as a solvent, it produces a large background of scattered neutrons and must be considered. There are two types of effects that must be considered in the scattering of neutrons from particles in a solution. They are called coherent and incoherent scattering. Coherent scattering arises from interference effects of a single neutron scattered from all of the nuclei in a given molecule and is very dependent on the scattering vector $\vec{q}$ (or scattering angle θ). Incoherent scattering can be thought of as the collective scattering of many neutrons from a single nucleus and is independent of the scattering angle θ. Incoherent scattering arises from fluctuations of scattering lengths within a given nucleus.

Even for a scatterer made up of a single isotope, not all of the scattering lengths

in Eq. 2.8 will be equal. This is because the scattering length of a nucleus depends on its spin state, and most isotopes have several spin states[28]. If we write the scattering length, b_j, as the sum of an average scattering length, $\bar{b}$, and a fluctuation scattering length, $\delta b_j = b_j - \bar{b}$, such that $< b_j >= \bar{b}$ and $< \delta b_j >= 0$, then we can separate Eq. 2.8 into two parts. (Here the angular brackets, $< . >$, denote a time or ensemble average over all isotopic spin states.) We find

$$\begin{aligned}
\frac{d^2\sigma}{d\Omega} &= \sum_{i=1}^{N}\sum_{j=1}^{N}(\bar{b}+\delta b_i)(\bar{b}+\delta b_j)e^{i\vec{q}\cdot\vec{r}_{ij}} \\
&= \bar{b}^2\sum_{i=1}^{N}\sum_{j=1}^{N}e^{i\vec{q}\cdot\vec{r}_{ij}} + \bar{b}\sum_{i=1}^{N}\sum_{j=1}^{N}\cancelto{0}{\delta b_i e^{i\vec{q}\cdot\vec{r}_{ij}}} + \bar{b}\sum_{i=1}^{N}\sum_{j=1}^{N}\cancelto{0}{\delta b_j e^{i\vec{q}\cdot\vec{r}_{ij}}} + \sum_{i=1}^{N}\sum_{j=1}^{N}\delta b_i\delta b_j e^{i\vec{q}\cdot\vec{r}_{ij}} \\
&= \bar{b}^2\sum_{i=1}^{N}\sum_{j=1}^{N}e^{i\vec{q}\cdot\vec{r}_{ij}} + \sum_{i=1}^{N}\delta b_i^2 + \sum_{i=2}^{N}\sum_{j=1}^{j<i}\cancelto{0}{\delta b_i\delta b_j e^{i\vec{q}\cdot\vec{r}_{ij}}} + \sum_{i=2}^{N}\sum_{j=1}^{j<i}\cancelto{0}{\delta b_i\delta b_j e^{-i\vec{q}\cdot\vec{r}_{ij}}} \\
&= \bar{b}^2\sum_{i=1}^{N}\sum_{j=1}^{N}e^{i\vec{q}\cdot\vec{r}_{ij}} + \sum_{i=1}^{N}\delta b_i^2
\end{aligned} \tag{2.9}$$

The first term in Eq. 2.9 is called coherent scattering and the second term is incoherent scattering. The middle two terms on the second line average to zero because $< \delta b_j >= 0$. The last two terms of the third line average to zero because scattering length fluctuations (or isotopic spin states) of different atomic nuclei within a molecule are uncorrelated.

In light scattering the polarizability, α, within a single molecule would take the place of the scattering length, b_j. Then, the second term in Eq. 2.9 would average to zero since the polarizability within a molecule is much more uniform (as seen by photons) than the neutron scattering length (as seen by neutrons). The reason for this is that, while a single nucleus may have different spin states, all atoms and

Tab. 2.1: Table of Neutron Scattering Length Density, Incoherent Scattering Cross Section and Index of Refraction

Material	$\rho = \frac{\bar{b}}{V}$	$\frac{4\pi(\bar{b^2}-\bar{b}^2)}{V}$	n
	$\times 10^{10}$ [cm^{-2}]^a	[cm^{-1}]^b	dimensionless
H_2O	-0.56	5.62	1.333^c
D_2O	6.3	0.14	1.329^c
Typical Protein	3.1	–	1.58^d
Deuterated Protein	8.5	–	–

[a]From Cantor and Schimmel[30] and Glatter and Kratky[31]

[b]Calculated from NCNR webpage: http://www.ncnr.nist.gov/resources/sldcalc.html

[c]Measured [d]From Hand et al.[32]

molecules of a homogeneous material have similar charge distributions and therefore have similar polarizabilities.

In SANS, one utilizes a bulk scattering length density, ρ, which for a given material, can be calculated by adding up the scattering lengths of the constituent nucleons and dividing by the volume of the particle, $\rho = \sum_j b_j/V$. Since our neutron scattering experiments are also in solution, we are interested only in the difference in scattering between the particle and solvent. The coherent scattered intensity is proportional to the square of the relative scattering length density, $(\Delta\rho)^2 = (\rho_p - \rho_s)^2$ where ρ_p is the average scattering length density of the particle and ρ_s is the average scattering length density of the solvent.

At low concentrations of particles, the incoherent scattered intensity is dependent mostly upon the scattering length density of the solvent. It is proportional to

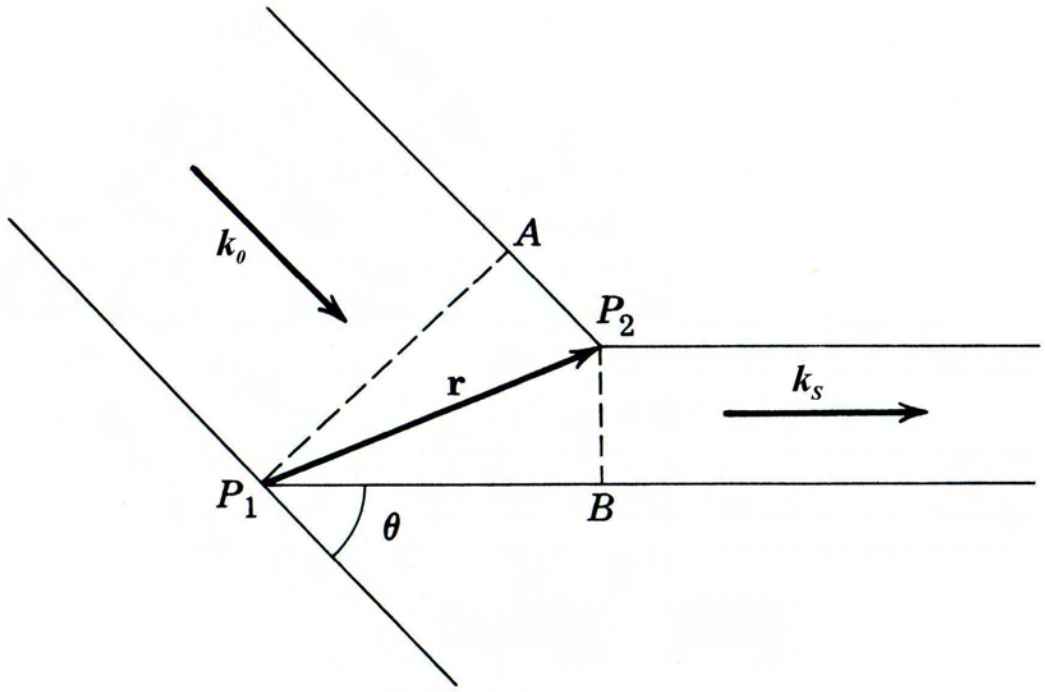

Fig. 2.4: "Diagram illustrating origin of phase difference in scattered waves. Coherent waves scattered from different positions have different phases associated with them arising from a difference in paths lengths traveled by the two scattered waves. Scattering from two identical scatterers located at P_1 and P_2." Figure and Captoion from *Physical Chemistry of Macromolecules*[26, Fig. 3-1(a)]

the mean squared scattering length of the solvent minus the square of the mean, $(\rho_{incoherent})^2 = \sum_j b_j^2 - (\sum_j b_j)^2/V = \sum_j \delta b_j^2/V$. Scattering length densities and indices of refraction for a few substances are shown in Table 2.1. This will be discussed more in Chapter 4.

2.2.4 Interference Between Waves Scattered from Different Positions Within a Particle Produce a Diffraction Pattern

The interactions between light and neutrons with matter are very different, but in both cases diffraction effects arise from interference between scattered coherent waves emanating from different points within the scattering volume. The theory for light scattering is very similar to that of neutrons with the exception of a few small changes.

Consider the waves originating from two scattering centers separated by position vector $\vec{r}$. A ray diagram of this is illustrated in Fig. 2.4. While the scattering amplitude is dependent on the type and strength of the interaction of the incident wave with the scatterer, the phase difference between coherent waves scattered from different positions in a scattering volume originates from the difference in path lengths traveled by each wave. This phase difference between the scattered waves can be expressed as

$$\Delta\phi = k_0 AP_2 - k_S P_1 B = \left(\vec{k_0} - \vec{k_S}\right) \cdot \vec{r} = \vec{q} \cdot \vec{r} \tag{2.10}$$

where $\vec{k_0}$ is the incident wave vector and $\vec{k_S}$ is the scattered wave vector. Here we have defined the scattering vector $\vec{q}$ as $\vec{k_0} - \vec{k_S}$. If the scattered wave is of the same wavelength as the incident wave, then the magnitude of the scattering vector, $q = |\vec{q}|$, can be written as $2k_0 \sin\frac{\theta}{2}$ where θ is the angle between the incident wave vector, $\vec{k_0}$, and the outgoing wave vector, $\vec{k_S}$. If the incident wave is coherent and the scattered waves are of similar amplitude, constructive interference will occur when $\Delta\phi$ is an

even multiple of π. Destructive interference will occur when $\Delta\phi$ is an odd multiple of π.

2.2.5 The Debye Equation

The scattering of radiation from a collection of scatterers of even very simple geometric shape is difficult to evaluate exactly. However, in the case where a very small fraction of the incident wave is scattered, one can use the first Born approximation to calculate the amplitude of the scattered wave with respect to the incident wave, as was done for both neutrons and photons in the previous section. In this approximation all scatterers see the same incident wave (in amplitude) and the relative phase difference between the scattered waves is dependent only on the path lengths as described above. This case was first studied by Debye[33], and was evaluated for a few simple geometric shapes such as spheres and random coils[34, 35, 36]. Many other simple geometric shapes and models such as rods, freely jointed chains, semi-flexible chains, and flexible star polymers have been investigated by others to study synthetic polymers in solution[31]. The general case for a collection of spheres, the Debye equation, will be derived in the following sections.

2.2.6 Scattering from Two Scatterers

Consider a monochromatic traveling plane wave as described by the expression $\Psi e^{i\vec{k}_0\cdot\vec{r}-i\omega t}$. If the wave is incident on two identical scatterers separated by a position vector, $\vec{r}$, (as shown in Fig. 2.4), the resulting scattered wave can be described by the sum of the waves scattered from each scatterer. (For simplicity, the scattered ampli-

tude will be assumed to be identical for each scatterer.) The scattering amplitude for a single scatterer will hereafter be denoted by Ψ_0.[4]

$$\begin{aligned}\Psi_S &= \Psi_0 e^{-i\omega t+i\vec{k}_0\cdot\vec{r}} + \Psi_0 e^{-i\omega t+i\vec{k}_S\cdot\vec{r}} \\ &= \Psi_0 e^{-i\omega t+i\vec{k}_0\cdot\vec{r}}(1+e^{i(\vec{k}_S-\vec{k}_0)\cdot\vec{r}}) \\ &= \Psi_0 e^{-i\omega t+i\vec{k}_0\cdot\vec{r}}(1+e^{i\vec{q}\cdot\vec{r}}) \end{aligned} \tag{2.11}$$

The reduced probability density (or equivalently the reduced differential scattering cross section) along a direction defined by $\vec{q}$, will be given by the norm squared of the scattered wave function, viz.,

$$\begin{aligned}\tilde{P}(\vec{q}) &= \frac{\Psi_S(\vec{q},t)\cdot\Psi_S^*(\vec{q},t)}{|\Psi_0|^2} \\ &= (1+e^{i\vec{q}\cdot\vec{r}})\cdot(1+e^{-i\vec{q}\cdot\vec{r}}) \\ &= (2+2\cos(\vec{q}\cdot\vec{r})) \\ &= 4\ \cos\left(\frac{\vec{q}\cdot\vec{r}}{2}\right)^2 \end{aligned} \tag{2.12}$$

2.2.7 Scattering from N Scatterers

The reduced scattering amplitude of a wave scattered by a molecule made up of N identical scatterers located at positions $\{\vec{r}_i\}$ can be expressed as a sum over scatterers, i, as

$$\begin{aligned}\tilde{f}(\vec{k}_0,\vec{q},\{\vec{r}_i\}) &= \frac{1}{\Psi_0}\sum_{i=1}^{N}\Psi_i \\ &= e^{i\phi_0}\sum_{i=1}^{N}e^{i\Delta\phi_{i0}}\end{aligned}$$

[4] According to the previous section $\Psi_0 = -k_0^2\alpha|E_0|\sin(\theta_1)\frac{e^{ik_0R}}{R}$ or $-b\frac{e^{ik_0R}}{R}$. Where $\vec{R}$ is the position of the detector with respect to the scatterer.

$$= \quad e^{i\vec{k_0}\cdot\vec{r_0}} \sum_{i=1}^{N} e^{i\vec{q}\cdot\vec{r}_{i0}} \tag{2.13}$$

where $\vec{r}_{i0}$ is the position vector of the i^{th} scattering center within the molecule, relative to some arbitrary origin within the molecule specified by $\vec{r}_0$ (say at the center of mass). The reduced probability density is the norm squared of the scattering amplitude. It is given by the double sum

$$\begin{aligned}
\tilde{P}(\vec{q}) &= |\tilde{f}(\vec{k}_0, \vec{q}, \{\vec{r}_i\})|^2 \\
&= f(\vec{k}_0, \vec{q}, \{\vec{r}_i\}) \cdot f^*(\vec{k}_0, \vec{q}, \{r_i\}) \\
&= \sum_{i=1}^{N}\sum_{j=1}^{N}(e^{i\vec{q}\cdot\vec{r}_i})(e^{-i\vec{q}\cdot\vec{r}_j}) \\
&= \sum_{i=1}^{N}\sum_{j=1}^{N} e^{i\vec{q}\cdot\vec{r}_{ij}}
\end{aligned} \tag{2.14}$$

where $\vec{r}_{ij} = \vec{r}_i - \vec{r}_j$ is defined as the relative distance between scatterers i and j.

If we define the particle distribution function $\rho(\vec{r}) = \frac{1}{V}\sum_{i=1}^{N}\delta^3(\vec{r} - \vec{r}_i)$ then we can rewrite the sum in Eq. 2.13 in an integral form:

$$\begin{aligned}
f(\vec{k}_0, \vec{q}, \{\vec{r}_i\}) &= e^{i\vec{k}_0\cdot\vec{r}_0} \sum_{i=1}^{N} e^{i\vec{q}\cdot\vec{r}_{i0}} \\
&= e^{i\vec{k}_0\cdot\vec{r}_0} \int_V d^3r \ \rho(\vec{r}) e^{i\vec{q}\cdot\vec{r}}
\end{aligned} \tag{2.15}$$

2.2.8 Solution Scattering: Averaging over Different Orientations

The molecules we are studying are suspended in solution rather than being located in solid crystals with a fixed orientation. They are randomly oriented. We can perform an angular average over all possible orientations of the molecule, thus removing any dependence of the scattered probability density $\tilde{P}$ on a specific direction of $\vec{q}$ with respect to the frame of reference of the molecule. This reduces our equation for

$\tilde{P}(\vec{q})$ to a simplified form with a dependence only on the *magnitude* of the scattering wave vector, q, (the angle between the incoming beam and the detector is implicit in this magnitude) and the relative distances between all scatterers within the molecule:

$$\begin{aligned}\tilde{P}(q) &= <\tilde{P}(\vec{q})>_{\Omega} \\ &= \sum_{i=1}^{N}\sum_{j=1}^{N}\int \frac{1}{4\pi} d\Omega e^{i\vec{q}\cdot\vec{r}_{ij}} \\ &= \sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\sin(qr_{ij})}{qr_{ij}}. \end{aligned} \tag{2.16}$$

2.2.9 Scattering from a Sphere

For scatterers distributed uniformly in a sphere of finite radius a with constant density $\frac{1}{V}$, we can integrate Eq. 2.15, vis.

$$\begin{aligned} f(\vec{k}_0,\vec{q},\vec{r_0}) &= e^{i\vec{k}_0\cdot\vec{r}_0}\frac{1}{V}\int_0^a r^2 dr \int d\Omega e^{i\vec{q}\cdot\vec{r}} \\ &= \frac{3(\sin(qa)-qa\cos(qa))}{q^3a^3}e^{i\vec{k}_0\cdot\vec{r}_0}. \end{aligned} \tag{2.17}$$

2.2.10 Scattering from N Spheres

Consequently if one integrates over N spheres of constant density $\frac{1}{V_i}$ and radius a_i centered at positions $\{\vec{r_i}\}$, then the scattering amplitude would be given by the sum

$$\begin{aligned} f(\vec{k}_0,\vec{q},\vec{r_0}) &= e^{i\vec{k}_0\cdot\vec{r}_0}\sum_{i=1}^{NM} e^{i\vec{q}\cdot\vec{r}_{i0}} \\ &= e^{i\vec{k}_0\cdot\vec{r}_0}\sum_{i=1}^{N} e^{i\vec{q}\cdot\vec{r}_{i0}}\sum_{j=1}^{M} e^{i\vec{q}\cdot\vec{r}_{ji}} \\ &= e^{i\vec{k}_0\cdot\vec{r}_0}\sum_{i=1}^{N} e^{i\vec{q}\cdot\vec{r}_{i0}}\frac{1}{V_i}\int_0^{a_i} r^2 dr \int d\Omega e^{i\vec{q}\cdot\vec{r}} \\ &= e^{i\vec{k}_0\cdot\vec{r}_0}\sum_{i=1}^{N}\frac{3(\sin(qa_i)-qa_i\cos(qa_i))}{q^3a_i^3}e^{i\vec{q}\cdot\vec{r}_{i0}}. \end{aligned} \tag{2.18}$$

Thus, the scattered intensity from a collection of spherical scatterers can be described by the equation

$$\tilde{P}(\vec{q}) = \sum_{i=1}^{N}\sum_{j=1}^{N}\left(\frac{3(\sin(qa_i)-qa_i\cos(qa_i))}{q^3a_i^3}\right)\left(\frac{3(\sin(qa_j)-qa_j\cos(qa_j))}{q^3a_j^3}\right)e^{i\vec{q}\cdot\vec{r}_{ij}}. \tag{2.19}$$

If we perform the angular averaging (as in Eq. 2.16) we get the resulting equation for a collection of spherical scatterers with fixed relative internal positions but an overall random orientation:

$$\tilde{P}(q) = \sum_{i=1}^{N}\sum_{j=1}^{N}\left(\frac{3(\sin(qa_i)-qa_i\cos(qa_i))}{q^3a_i^3}\right)\left(\frac{3(\sin(qa_j)-qa_j\cos(qa_j))}{q^3a_j^3}\right)\frac{\sin(qr_{ij})}{qr_{ij}}. \tag{2.20}$$

This equation was first explored by Debye[33, 31] and is used later in this thesis to calculate the expected scattering of neutrons from simple bead models of clathrin. It assumes that the molecule is made up of scatterers having identical scattering length density relative to the solvent.

2.2.11 The Guinier Approximation

Consider a rotationally averaged probablility density $P(q)$ (Eq.2.16) for a molecule with a limited extent, R, such that $r_{i0} < R\ \forall i$. If we look only at small q, where $qR \ll 1$ (That is, if we are only looking at small scattering angles or using a large wavelength in comparison to R), we can expand the function $\sin(x)$ in a Taylor's expansion $\sin(x) = \sum_{i=0}^{\infty}\frac{x^{2i+1}}{2i+1!} = x - \frac{x^3}{6} + \frac{x^5}{120} - O(x^7)$.

By inserting this expression in Eq. 2.16, one gets

$$\begin{aligned}
\tilde{P}(q) &= \sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\sin(qr_{ij})}{qr_{ij}} \\
&= \sum_{i=1}^{N}\sum_{j=1}^{N}1-\frac{(qr_{ij})^2}{6}-O(qr_{ij})^4 \\
&\approx N^2\left(1-\frac{(qR_g)^2}{3}+\frac{(qR_g)^4}{18}-O(qR_g)^6\right) \\
&\approx N^2 e^{-\frac{1}{3}(qR_g)^2}
\end{aligned} \tag{2.21}$$

In this last step we have used the identity $R_g^2 = \frac{1}{2N^2}\sum_{i=0}^{N}\sum_{j=0}^{N} r_{ij}^2$. This approximation from small angle scattering was derived by Guinier[37] and will be used later to determine the radius of gyration, R_g from static light scattering and neutron scattering experiments performed on clathrin.

The Guinier approximation[26, p. 305] is valid only when $\frac{1}{20} < R_g q < \frac{1}{2}$. The upper limit on R_g arises because of the expansion of $\sin(R_g q)$ used to derive the equation. The lower limit on R_g is due to the accuracy with which the scattered light intensity can be measured, and therefore depends upon the sensitivity of the experimental apparatus that is used. In order to give a 10% - 50% drop in scattered intensity by variation of the detector angle, $\frac{R_g}{\lambda}$ must be greater than $\frac{1}{20}-\frac{1}{10}$. In my static light scattering (SLS) experiments, the drop in intensity observed was 8% in the q range that was used to determine the measured R_g. Prior to each experiment, the instrument was calibrated to within 1% for all angles used in the measurement.[5]

[5] The equation for the lower limit on particle size is $\frac{R_g}{\lambda} > \frac{1}{4\pi}\sqrt{3\ln\left(\frac{I(\theta=180)}{I(\theta=0)}\right)}$.

2.3 Dynamic Light Scattering

In the previous section we ignored the time dependence of the scattered wave as well as any motion or dynamics of the molecule from which we were scattering. Suppose that, instead of calculating an instantaneous probability density, which was done implicitly, we included the time dependence of the scattered probability density using time dependent scattering positions $\{r_i(t)\}$ instead of static ones. Our number density would then become time dependent $\rho(\vec{r},t) = \frac{1}{V}\sum_{i=1}^{N}\delta^3(\vec{r}-\vec{r}_i(t))$ and, following Eq. 2.15 and now including a time dependent factor $e^{-i\omega_0 t}$, the scattered wave would be

$$\Psi(\vec{k}_0,\vec{q},\{r_i(t)\},t) = \Psi_0 e^{i\vec{k}_0\cdot\vec{r}_0 - i\omega_0 t}\int_V d^3r\rho(\vec{r},t)e^{i\vec{q}\cdot\vec{r}}. \tag{2.22}$$

If we wanted to proceed further, then we would have to know how $\{r_i(t)\}$ evolves in time. For simplicity, consider the case of N particles whose initial positions at time, $t = 0$, are $\vec{r}_0$ and whose motion is determined by diffusion. In that case, the average position of the particle would be $\vec{r}_0$ and the mean squared displacement, $< r^2 >$, would be $6Dt$. We could then write the time dependent number density as $\rho(\vec{r},t) = \frac{N}{\left(\sqrt{4\pi Dt}\right)^3}e^{-\frac{(\vec{r}-\vec{r}_0)^2}{4Dt}}$. And the time dependent scattered wave from N particles undergoing diffusive motion could be written as

$$\begin{aligned}\Psi(\vec{k}_0,\vec{q},r_0,t) &= N\Psi_0 e^{i\vec{k}_0\cdot\vec{r}_0 - i\omega_0 t}\int_V d^3r\frac{e^{-\frac{r^2}{4Dt}}}{(\sqrt{4\pi Dt})^3}e^{i\vec{q}\cdot\vec{r}} \\ &= N\Psi_0 e^{i\vec{k}_0\cdot\vec{r}_0 - i\omega_0 t}e^{-Dq^2t}.\end{aligned} \tag{2.23}$$

Using the above equation we can calculate the scattered intensity autocorrelation

function

$$\begin{aligned} c(\tau) &= <\tilde{P}(t)\tilde{P}(t+\tau)> \\ &= \frac{N^4}{|\Psi_0|^4}\int_0^\infty dt\ \Psi^*(\vec{k}_0,\vec{q},r_0,t)\Psi(\vec{k}_0,\vec{q},r_0,t) \\ &\qquad \times \Psi^*(\vec{k}_0,\vec{q},r_0,t+\tau)\Psi(\vec{k}_0,\vec{q},r_0,t+\tau) \\ &= \frac{N^4}{4Dq^2}e^{-2Dq^2\tau}. \end{aligned} \tag{2.24}$$

This equation is the basis for dynamic light scattering as used later to determine the diffusion coefficient of clathrin triskelia in solution.

2.4 Hydrodynamic Theory of Macromolecules

Hydrodynamic techniques play a major role in the characterization of macromolecules. Experimental hydrodynamics measurements include viscometric, sedimentation and diffusion measurements. However, relating these measurements to desired structural information requires a theoretical formalism enabling the calculation of hydrodynamic properties from a model of the macromolecule under study.

In the case of flexible and semiflexible linear chains, theoretical work was initiated by Kirkwood and Riseman[38]. Hydrodynamic formalisms for more complicated rigid structures were available only for ellipsoids and cylinders. Early theoretical work on the hydrodynamics of rigid particles with arbitrary shape began with the work of Bloomfield et al.[39, 40]. In this work the theory of Oseen, Burgers, Kirkwood and Reisman[41, 42, 38] were used to calculate hydrodynamic properties of proteins (hemocyanin, phycocyanin, and fibrinogen) and viruses (Tobacco mosaic virus, T2 bacteriophage, and λ phage). Dynamic light scattering and sedimentation

measurements on each example were carefully compared with theoretical estimates from multi-subunit bead models with dimensions and geometry in accordance with electron microscope images that had been taken by other researchers.

Advances to the work of Oseen were later made by Rotne and Prager[43] as well as Yamakawa[44] in the form of a modified hydrodynamic interaction tensor. Later with the aid of modern computers, Bloomfield and Garcia de la Torre improved upon Kirkwood and Reisman's approximation[45, 46, 47]. X-ray structures of proteins as well as cryo-EM reconstructions of macromolecular assemblies have dramatically improved the detail of structural models of some biological macromolecules. As an example of how the field has progressed, fluorescence correlation spectroscopy has recently been used to study tubulin rings, showing remarkable agreement with the predicted hydrodynamic properties using an atomic model of a tubulin ring structure that was determined by combining cryo-EM and X-ray diffraction measurements[48].

2.4.1 Inversion of the Hydrodynamic Equations (HYDRO)

In 1927, Oseen showed that the effect of a point source of friction on the flow of a Newtonian fluid of viscosity, η, was to reduce the flow velocity, $\vec{v}_0$, at a position, $\vec{r}$, by an amount, $\delta\vec{v} \equiv \vec{v} - \vec{v}_0$ where $\vec{v}_0$ is the velocity of flow in the absence of the force and $\vec{v}$ is in the presence of the force. $\delta\vec{v}$ is determined by the product of a hydrodynamic interaction tensor, $\mathbf{T}$, with the applied force $\vec{F}$,

$$\delta\vec{v} = -\mathbf{T}\cdot\vec{F} = -\frac{1}{8\pi\eta r}\left(\mathbf{I} + \frac{\vec{r}^{T}\vec{r}}{r^2}\right)\cdot\vec{F}. \tag{2.25}$$

A particle moving with velocity $\vec{v}$ through a stationary fluid of viscosity, η,

produces a force, $\vec{F}$, proportional to $\vec{v}$ on the surrounding fluid in the direction of motion, $\vec{F} = \zeta\vec{v}$ where ζ is the coefficient of friction. Neighboring particles will therefore reduce the flow velocity and the viscous force between the surrounding fluid and nearby particles. This is the basic principle behind hydrodynamic modeling of rigid biological macromolecules.

The viscous force on a particle can be approximated by taking into account the effect of forces by neighboring particles on the flow velocity of the fluid using Oseen's tensor,

$$\vec{F} \;=\; \zeta(\vec{v} - \mathbf{T}\cdot\vec{F}) \tag{2.26}$$

from which one finds

$$\mathbf{Q}\cdot\vec{F} = (\mathbf{I} + \zeta\mathbf{T})\cdot\vec{F} = \zeta\vec{v} \tag{2.27}$$

where $\mathbf{T}$ is a sum over all neighboring particles. Setting the force on the right hand side of the Eq. 2.26 to $\vec{F} \approx \zeta\vec{v}$, we get an approximation generally referred to as the Kirkwood-Riseman approximation[38]. If we solve the system of equations by inverting the matrix $\mathbf{Q} = \mathbf{I} + \zeta\mathbf{T}$ for each element in the particle then we can calculate the entire friction tensor of a fixed collection of particles. This is the principle behind current work involving the description of the hydrodynamics of polymers and biological macromolecules.

Zwanzig[49] showed that if Oseens' hydrodynamic interaction tensor (defined in Eq. 2.25) is used to calculate the friction of a bead model of finite size, then $\mathbf{Q}$ is not necessarily positive definite and can, in some cases, give negative friction coefficients. Rotne and Prager[43] and Yamakawa[44] corrected this by including

the effect of particle size in the hydrodynamic interaction tensor of Eq. 2.25. This was done by using Stokes' law for the coefficient of friction of a sphere of radius a, $\zeta = 6\pi\eta a$ for overlapping and nonoverlapping spheres. The nonoverlapping case was later generalized by Garcia de la Torre et al.[45] for spheres of different sizes

$$\mathbf{T}(a_i, a_j, \vec{r}_{ij}) = \frac{1}{8\pi\eta r_{ij}} \left[\mathbf{I} + \frac{\vec{r}_{ij}^T \vec{r}_{ij}}{r_{ij}^2} + \frac{a_i^2 + a_j^2}{r_{ij}^2} \left(\frac{1}{3}\mathbf{I} - \frac{\vec{r}_{ij}^T \vec{r}_{ij}}{r_{ij}^2} \right) \right] \tag{2.28}$$

the overlapping case ($r \leq 2a$) is

$$\mathbf{T}(a, \vec{r}_{ij}) = \frac{1}{6\pi\eta a} \left[\left(1 - \frac{9}{32}\frac{r_{ij}}{a} \right) \mathbf{I} + \frac{3}{32} \frac{\vec{r}_{ij}^T \vec{r}_{ij}}{a r_{ij}} \right] \tag{2.29}$$

The remaining case, calculations with overlapping spheres of different sizes, a_i and a_j, use Eq. 2.29 with a bead radius of$a = \frac{1}{2}(a_i^3 + a_j^3)^{\frac{1}{3}}$ so that the volume of the two beads of equal size is the same as the volume of bead i and j. See Carrasco et al. for further discussion[50].

HYDRO[6] is a program written by Garcia de la Torre et al.[51]. Using the theory outlined above, HYDRO calculates the friction coefficient of rigid macromolecules of arbitrary shape that are represented by bead models. I constructed an ad hoc bead model for clathrin based on EM, cryo-EM, and crystal structure measurements of clathrin arms. This model, as well as the results of HYDRO calculations performed with it, are described in Chapter 3.

The above described treatment for the direct inversion of the hydrodynamic equations represents a system of linear equations of 3 unknowns for each element in the model. This is accomplished by formulating the equations into a $3N \times 3N$ super

[6] http://leonardo.fcu.um.es/macromol/

matrix $\mathbf{Q_N}$. This matrix is then inverted by Gauss-Seidel iteration[51]. The Gauss-Seidel method is a fast iterative matrix inversion technique similar to the Gauss-Jordan method. Computation time scales approximately as N^3 and necessary storage space grows as $9N^2$. This limits practical models to $N \sim 2000$ elements on modern computers in 2006($N \sim 100$ in 1981). This limitation is mainly due to the speed at which a computer can invert a matrix of size $(3N) \times (3N)$, i.e. the rate of convergence of the Gauss-Seidel algorithm[52, 53].

2.4.2 The Analogy between Electrostatics and Hydrodynamics (ZENO)

It is possible to calculate a rotationally averaged translational friction coefficient for a rigid object using an alternative method proposed by Hubbard and Douglas[54]. If one performs an angular average of Oseen's hydrodynamic interaction tensor, (Eq 2.25), it becomes a simple scalar.

$$\begin{aligned} < \mathbf{T} >_\Omega &= \frac{1}{4\pi} \int d\Omega \ \mathbf{T} \\ &= \frac{1}{4\pi} \int d\Omega \frac{1}{8\pi\eta r} \left(\mathbf{I} + \frac{\vec{r}^T \vec{r}}{r^2} \right) \\ &= \frac{1}{8\pi\eta r} \left(\mathbf{I} + \frac{1}{3}\mathbf{I} \right) \\ &= \frac{1}{6\pi\eta r} \mathbf{I} \end{aligned} \tag{2.30}$$

With this angular averaged hydrodynamic interaction tensor, it can then be shown that the Navier-Stokes equations reduce to an equivalent electrostatics problem where the angular averaged translational friction coefficient of a particle can be related to the capacitance, C, on its surface as if it were conducting[54].

$$< f >_{\Omega} = 6\pi\eta C. \tag{2.31}$$

The capacitance, C, is easier to calculate for complicated objects than is the matrix inversion described in the section above, and has a less restrictive limit to the number of particles used in the model. Also, this calculation requires only that a surface be defined in some way and is not, in principle, reliant on bead modeling. The capacitance calculation is done by a numerical path integration technique[55] in the program ZENO[7].

I used ZENO for calculations performed on high resolution triskelion models as well as on complete clathrin baskets. In some cases as many as 7000 elements were used. For models where HYDRO and ZENO could both be used, the rotationally averaged, translational friction coefficients agreed in all cases tested to within 1%.

[7] http://www.stevens.edu/zeno/

3. LIGHT SCATTERING FROM A CLATHRIN TRISKELION IN SOLUTION

3.1 Introduction

Triskelia assemble at low pH to form baskets [14] or, at higher pH, in the presence of clathrin assembly proteins (e.g. AP2, AP180, etc.)[56]. It is not currently known why these conditions favor assembly. Possible mechanisms supporting assembly at low pH include conformational changes in the tertiary structure of the triskelia, which might relieve steric hindrances to basket assembly, and enhanced interactions between the intertwined triskelia that form the basket struts. The role of AP's may be to increase mechanical linkages between the triskelia [56, 57, 58].

In this chapter I report light scattering measurements on clathrin in solution, which include conditions where assembly occurs. By static light scattering (SLS) I determine the radius of gyration (R_g), and by dynamic light scattering (DLS) I determine the hydrodynamic radius (R_H), of triskelia that have not polymerized into baskets. Although one does not obtain atomic and molecular coordinates as one might from a diffraction experiment or cryo-EM tomography, light scattering provides weighted averages that are sensitive to different aspects of the spatial distribution of

triskelion mass. In addition, the light scattering techniques are non-invasive and thus allow one to assess the native state of clathrin in solution.

In this chapter I will compare the measured values of R_g and R_H with corresponding quantities calculated for a simplified model of a clathrin triskelion. The model is designed, in particular, to show dependences on triskelial pucker and leg bending. Where possible, I use leg segment dimensions based on X-ray and EM structural data. Using the model, I infer that both R_g and R_H are intimately linked to the geometric attributes of the triskelia and that, taken together, they provide a sensitive measure of the solution conformations of these complex objects. Analysis of the measurements indicates that, over a large range of pH, individual triskelia are puckered in solution.

Recently, 3D structures of "D6 barrel" clathrin cages have been used to infer a molecular model of a single clathrin triskelion [16]. The D6 barrel is a polyhedral assembly of 36 clathrin triskelia and associated AP2 assembly complexes. Calculations of R_g and R_H based on this model agree fairly well with our light scattering measurements on free clathrin triskelia, but indicate that a triskelion assumes a slightly more compact structure when incorporated into the cages.

3.2 Experimental Methods

3.2.1 Clathrin Purification

Clathrin was purified from coated vesicles of bovine brain as previously described by Morgan et al. [57]. Here I would like to thank and acknowledge K. Prasad and E. Lafer

of the University of Texas Health Science Center in San Antonio, who performed all of the protein purification of clathrin. Briefly, clathrin coated vesicles were separated from homogenized bovine brain tissue by a two step density gradient centrifugation in low concentrations of sucrose and D_2O according to Nandi et al.[59]. The purified coated vesicles are then dialyzed in disassociation buffer (10mM Tris-HCl, pH 8.5) for 15 h. This causes the clathrin to uncoat the vesicles. Once uncoated the vesicles are then pelleted by centrifugation for 1 hr at 200,000 x g. The supernatant containing clathrin and associated proteins is then dialyzed in 0.5 M Tris-HCl, pH 8.0, and gel filtered on a Sephacryl S-300 column to separate clathrin, which elutes first, separate from the smaller molecular weight associated proteins.[60] Once purified the protein was stored in 0.5 M Tris HCl pH 7.0, 3mM DTT, at a protein concentration of 3 to 4 mg/mL at 4° C, and used within 2 weeks.

Every 24 hrs, prior to experiments, the protein was clarified[1] at 400,000 x g for 15 minutes to remove any protein that was no longer viable due to misfolding or proteolysis. At the end of 2 weeks the amount of protein removed in this process was close to half of the total protein at which point the sample was judged to be no longer useful for experiments.

3.2.2 Sample Preparation

For light scattering, buffers were filtered three times through 0.22 μm cellulose acetate filters (Corning) prior to use. Buffer was exchanged by overnight dialysis through a

[1] Clarification simply means to sediment out particulates by centrifugation, taking only the top of the solution and leaving behind a solid pellet.

10,000 Da molecular weight cutoff membrane (Pierce) into 10mM Tris HCl pH 8.0, 3mM DTT. The protein was diluted to 0.34 mg/mL and clarified by centrifugation at 400,000 x g for 15 minutes in an Optima Max Ultracentrifuge (Beckman), using an MLA130 rotor. After centrifugation, the entire supernatant was taken. The protein concentration was found to remain constant to within 0.01 mg/ml.

Soluble triskelia were prepared for light scattering at different pH by adding 100 μL of 1M MES buffer to 0.9 mL of clarified protein solution. Samples were then incubated one hour on ice and centrifuged at 400,000 x g for 15 minutes to remove any assembled baskets. Only the top 80% of the supernatant was carefully taken[57]. The preparation was confirmed, by dynamic light scattering, to consist of mono-disperse clathrin triskelia. The clathrin concentration in these samples was determined by absorbance measurements at a wavelength of 280 nm, using an absorption coefficient, $\epsilon_{280} = 6.94 \times 10^3\ M^{-1}cm^{-1}(1.07g^{-1}cm^{-1})$, as estimated from the amino acid sequence from *Bos Taurus* and following the calculation given by Reference [61].

3.2.3 Light Scattering Instrument

We used a Brookhaven light scattering instrument (Brookhaven Instruments Corp., Holtsville, NY) for both static and dynamic light scattering measurements. An Argon-ion laser (Lexel) beam, emitting at a wavelength of 514.5 nm, was directed onto the sample. Depending on the sample concentration, the incident intensity was adjusted to give a detectable scattered intensity of 100,000 counts per second at 90° scattering angle and fixed detection pinhole (400 μm or 1 mm in diameter). The scattered light was collected by a photomultiplier tube (EMI-PMT Model 9863) and processed by a

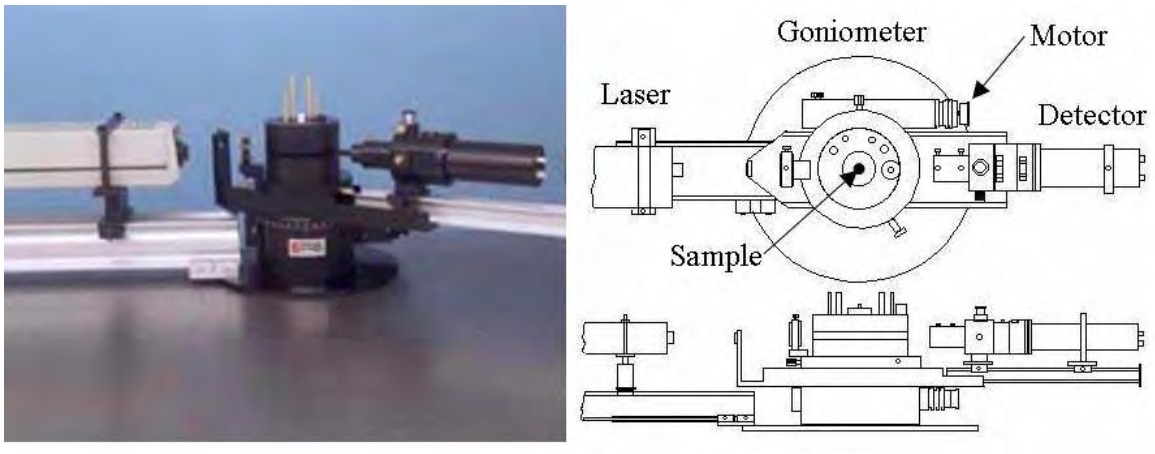

Fig. 3.1: Picture and diagram of light scattering instrument used. The beam from an argon ion laser was focused onto the sample through a steering lense assembly. The detector was mounted to an adjustable goniometer which could be rotated to the desired angle by a computer controlled motor. The detector signal was the processed by a digital correlator. Pictures and diagrams are take from Brookhaven Instruments Corp. ©website: http://www.bic.com/BI-200SMls.htm[62]

BI-9000 AT data acquisition board (Brookhaven Instruments Corp.). For DLS, the board calculates and generates the intensity-intensity time-correlation functions.

A 1 mL volume of sample, loaded in a 1 inch diameter cylindrical light scattering cuvette, was placed in a decalin index-matching bath attached to a precision goniometer (Brookhaven Instruments). The temperature of the bath was measured before and after each measurement, using a digital thermometer ($\pm 0.1^{\circ}$C).

3.2.4 Static Light Scattering

Scattered light intensity was measured every 2°, from a scattering angle of $\theta = 50^{\circ}$ to 150°. Background scattering intensity from the cuvette and solvent was less than 1% of the detected signal in all experiments. In the Guinier regime ($qR_g << 1$), the angular dependence of the scattered intensity, I(q), is well approximated by a Gaussian profile with a width related to the radius of gyration,

$$I(q) \approx \exp\left(-\frac{1}{3}q^2 R_g^2\right) \tag{3.1}$$

where the scalar quantity q is the magnitude of the Bragg scattering wave vector [63], viz.,

$$q = \frac{4\pi n}{\lambda}\sin\left(\theta/2\right) \tag{3.2}$$

λ, θ and n being the wavelength, the scattering angle, and the refractive index of the solution, respectively. The logarithm of the intensity profile was plotted against q^2 and fitted by a straight line[26], using the method of least squares (unweighted). The slope of the line is $1/3\ R_g^2$. For any given sample, the measurements were repeated four times and the fits yielded values for R_g reproducible within 3% (0.8 nm).

3.2.5 Dynamic Light Scattering

Dynamic light scattering was performed at scattering angles of 90° and 150°. Photon counts collected over prescribed intervals were auto-correlated over a time range, τ, to provide the autocorrelation function of the scattered light, $c(\tau)$. Data were collected in the homodyne mode, from which the field correlation function, $g(\tau)$, was determined by the following formula:

$$c(\tau) = A\ \left(1 + B\ |g(\tau)|^2\right) \tag{3.3}$$

where A and B are constants [63].

In the homodyne mode only the scattered light impinges on the detector. This is in contrast to the heterodyne mode where a small portion of the incident beam is mixed with the scattered light on the detector[2]. In homodyne mode the light intensity autocorrelation function is proportional to the square of the field correlation function, giving a relaxation timescale, $\tau \propto 2Dq^2$, in accord with Eq. 2.24.

Using the cumulant method, I fitted the logarithm of the measured field correlation function up to the 2nd cumulant by the equation [64]:

$$\frac{1}{2}\ln|g(\tau)^2| = \ln|g(\tau)| = -Dq^2\tau + \frac{1}{2}\mu q^4\tau^2 \tag{3.4}$$

where q is given by Eq. 3.2, D is the average translational diffusion coefficient, and μ is related to the width of the distribution of the diffusion coefficient. Both D and

[2] The heterodyne mode is a technique by which the intensity autocorrelation function recorded in an experiment is directly proportional to the field correlation function, $g(\tau)$, and decays with a relaxation timescale, $\tau \propto Dq^2$.

μ are the fitting parameters. I then determined the mean hydrodynamic radius, R_H, from the Stokes-Einstein relation:

$$D = k_B T / 6\pi\eta R_H \tag{3.5}$$

where k_B is Boltzmann's constant, T the temperature, and η the solvent viscosity. For a given sample, two correlation functions were collected at $\theta = 90^\circ$ and two at $\theta = 150^\circ$ and fitted with the expression in Eq. 3.5, yielding values for R_H reproducible within 1% (0.2 nm). Here the buffer viscosity (for 0.1 M MES) was taken to be 1.05 times that of water, in accordance with measurements made with an Ostwald viscometer at 20 °C (data not shown).

3.2.6 Hydrodynamic Modeling

Transport properties of clathrin triskelia were calculated using HYDRO [51] and ZENO [55]. To employ these programs, triskelia of differing shapes were approximated by groups of connected spheres. Calculations primarily were performed with the program HYDRO [53, 51], which uses a hydrodynamic interaction tensor and Stokes' solution for flow around a sphere to determine the frictional force on a rigid bead model by an iterative matrix inversion of the hydrodynamic equations [53]. The resultant frictional coefficient of the structure then is related to the hydrodynamic radius, R_H. A recent experimental study of the hydrodynamic properties of nanoscopic rings, formed from the cytoskeletal protein tubulin, showed extremely good quantitative agreement between measurements and parameters computed by this method [48]. ZENO uses a Monte Carlo, numerical path integral technique to estimate the

translational diffusivity by making an analogy between isotropic angular averaged hydrodynamics and electrostatics [54, 55]. Like HYDRO, it also uses an iterative algorithm. It allows for more complex models than does HYDRO, the latter being limited to 2000 beads. When I compared these programs, ZENO gave results for R_H within 1% of those determined by HYDRO, for all models tested.

3.3 Results

3.3.1 Light Scattering

Clathrin assembles into closed, basket-like structures at low pH and low salt concentration [14]. The baskets can be removed by centrifugation, while unassembled clathrin remains in the supernatant. Above a pH of 6.5, basket formation occurs relatively slowly and is only weakly dependent on pH, whereas at pH 6.0 most of the clathrin is quickly incorporated into closed polyhedra. After centrifuging the baskets from these samples, DLS measurements made on the supernatants show highly mono-disperse scatterers having an average value of $r_H = 16.9$ nm for the apparent hydrodynamic radius. The reduced second moment, μ/D^2 had an average value of 0.06 ± 0.03 and did not vary significantly with pH. Typical results from light scattering measurements are shown in Fig. 3.2 and 3.3.

Fig. 3.2 shows typical SLS data for clathrin in solution, where the vertical lines indicate the q range over which the data were fitted to a linear function to determine the value of R_g (see Eq. 3.1). In Fig. 3.3 typical DLS data are shown as well as the result of the quadratic fit (see Eq. 3.4) used to determine the value of R_H (see Eq.

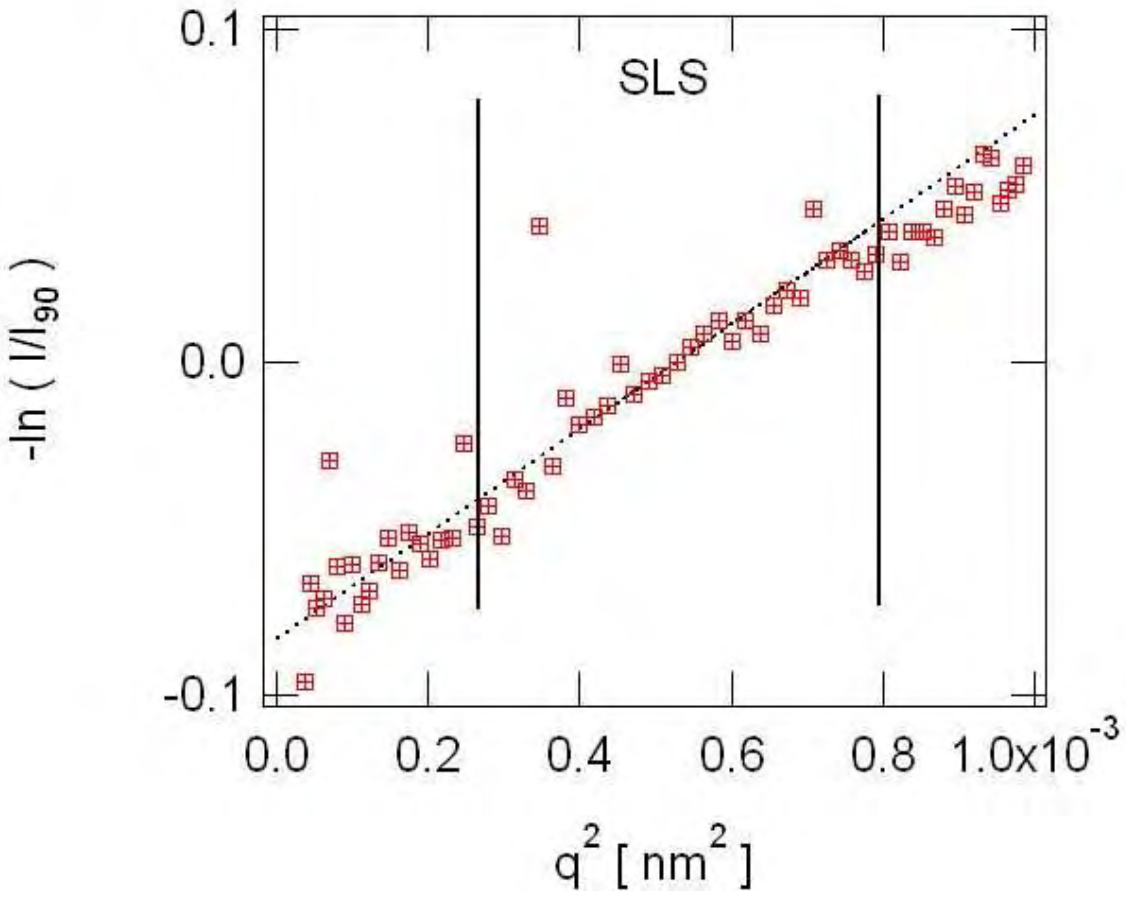

Fig. 3.2: Plotting $\ln I(q)$ vs. q^2 yields a slope of $\frac{1}{3}R_g^2$. Typical static light scattering data, for clathrin in solution at pH 6.5. The open circles indicate static light scattering data. The dotted line indicates a linear fit to the data from q^2=0.25 to 0.80 x10^3 nm^{-2} ($\theta = 60°$ to 120°). [62]

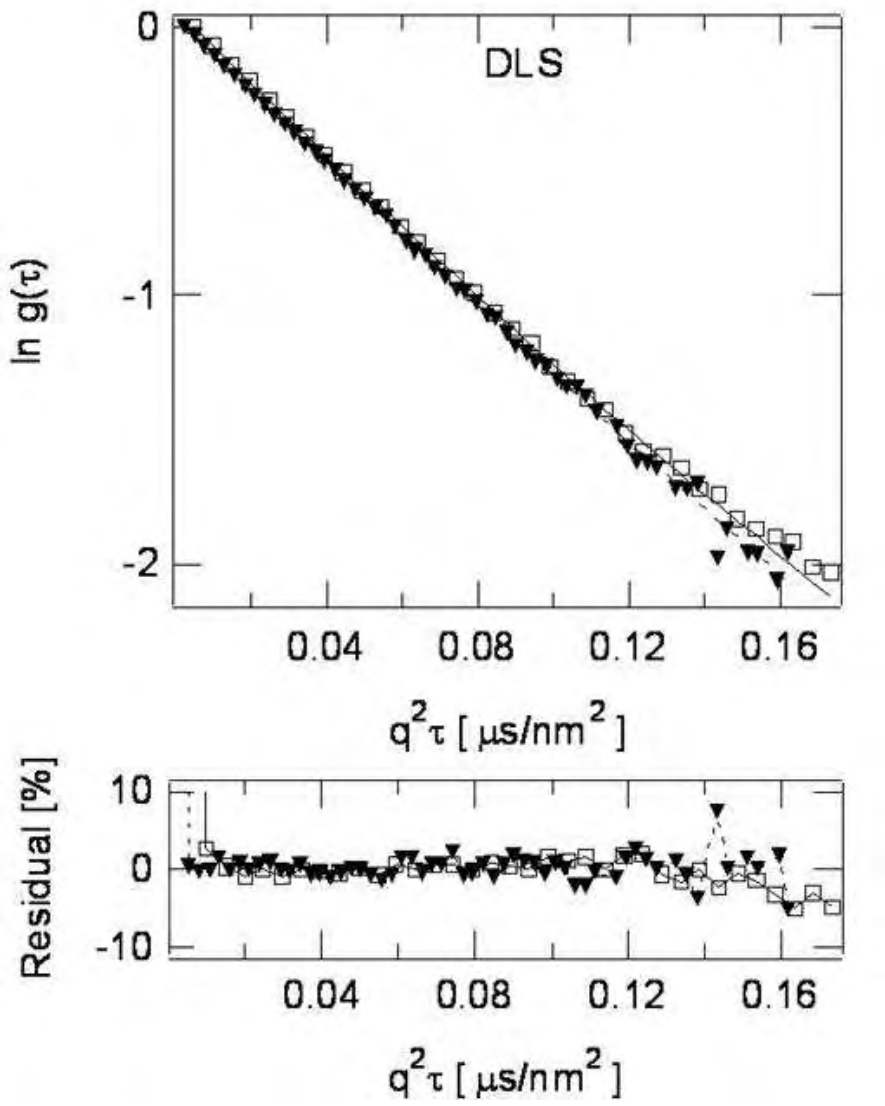

Fig. 3.3: Plotting $\ln g(\tau)$ vs. $q^2\tau$ gives a slope of $-D$. (top) Typical dynamic light scattering data, for clathrin in solution at pH 6.5, taken at angles of 90° and 150°. Open squares indicate dynamic light scattering data taken at an angle of 150°. Solid triangles indicate light scattering data taken at an angle of 90°. The solid and dotted lines indicate quadratic (in τ) fits to the data (see Eq. 3.4)[62]. (bottom) Residuals of quadratic fits to the dynamic light scattering data shown in % deviation. Typical points on fitted curves lie within 1% of the experimental data.

3.4 and 3.5). Fig. 3.4 shows results derived from light scattering measurements as described above.

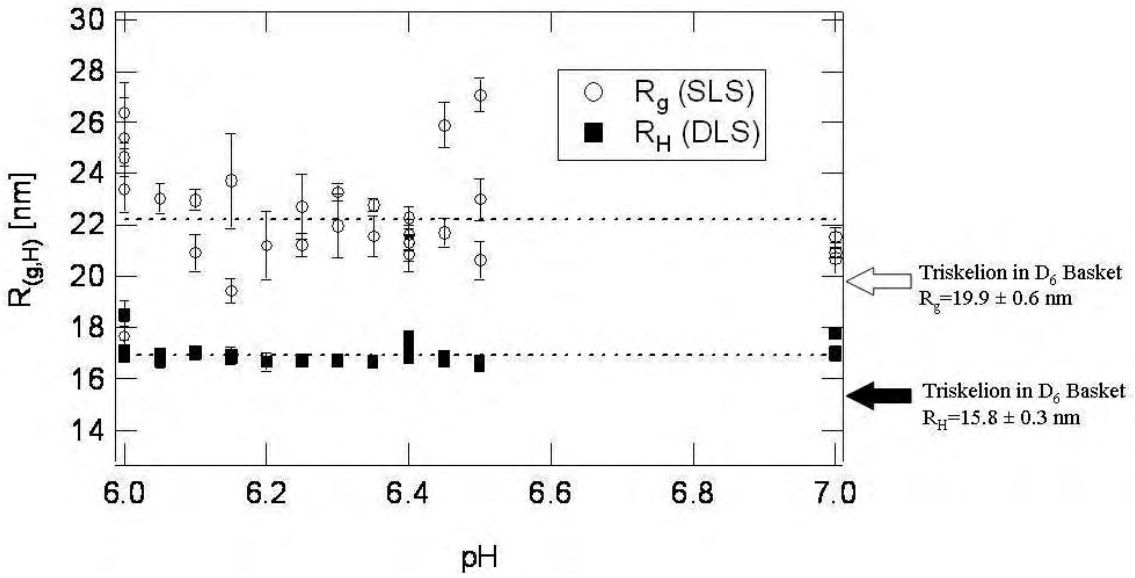

Fig. 3.4: Experimental R_H and R_g vs. buffer pH. Radius of gyration (R_g) and hydrodynamic radius (R_H) of free clathrin triskelia measured by static light scattering (SLS) and dynamic light scattering (DLS). The abscissa indicates the pH of the buffered solution. Both measurements appear to be independent of pH. The arrows point to the values calculated by ZENO [54, 55] for the triskelial structure noted for clathrin in an assembled basket[16]. Each data point was taken on a different sample. Error bars represent standard deviations of four measurements on the same sample. [62]

We measured a hydrodynamic radius, R_H, and radius of gyration, R_g, for free clathrin in solutions of different pH and found that, within experimental error, they are almost independent of pH. As previously mentioned, clathrin baskets were re-

moved by centrifugation prior to light scattering. Even so, at low pH clathrin continues to assemble, albeit at slower rates due to the reduced clathrin concentrations resulting from the elimination of polymerized material. (The protein concentration after centrifugation at pH 6.0 was 0.11±0.03 mg/ml). Thus, data were acquired as soon as possible after the centrifugation step. Measurements took 2 to 10 minutes to complete. Ignoring the data taken at pH 6.0, which were the most difficult points to repeat because of the tendency of triskelia to self-assemble, one finds an average value of $R_g = 22.2 \pm 1.8$ nm and $R_H = 16.9 \pm 0.3$ nm, with no significant pH dependence. Light scattering measurements of triskelia immersed in pH 7.0, 0.5 M Tris buffer yield similar values (See Table 3.1 and 4.1).

The bold arrows in Fig. 3.4 indicate the values of R_H and R_g calculated by applying ZENO [54, 55] to the triskelial structure found in a D6 basket assembled in the presence of AP2 assembly complexes (PDB code 1XI4[16]) mentioned in the introduction. Interestingly, the R_H and R_g derived from the solution measurements are close to, but somewhat greater than, values calculated for the configuration of the clathrin triskelion in the basket (see Table 3.1).

3.3.2 Calculations Using the Cryo-EM Structure of a Triskelion

The calculations of R_H and R_g for the shape of a triskelion within a basket were initially performed with the program HYDRO. A modified 2.1 nm resolution structure for a clathrin triskelion in a D6 basket was generously provided by T. Kirchhausen (triskelion.pdb[5]). These coordinates were determined by combining partial structures determined by X-ray diffraction studies with cryo-EM tomography

of a clathrin basket[65]. This structure was used to generate a rigid bead model for a clathrin triskelion which the program HYDRO could then use to calculate an approximate R_H and R_g. The 192 coordinates of this structure located the centers of contiguous beads, these centers being, on average, 0.76 $\pm$ 0.10 nm apart. the bead diameter was chosen to equal the average leg thickness of 3.0 nm, determined from the partial crystal structure for the α-solenoid clathrin proximal leg(PDB code 1B89[22]). (See Fig. 2.2a.) In order to account for the β-propeller N-terminus, which was not resolved by the cryo-EM tomography, this model was completed by the addition of 5.0 nm beads to the ends of the legs in accordance with the partial crystal structure (PDB code 1BPO[11]), shown in Fig. 2.2b. The bead model described above for an entire clathrin triskelion is shown in Fig. 3.10e.

The radius of gyration, R_g, was calculated with this model, but a 0.3 nm hydration layer was added for the calculation of the hydrodynamic radius. (A single molecule of water is approximately 0.3 nm.) The results of HYDRO calculations on this structure are shown in the second and third rows of Table 3.1.

Recently, a higher resolution cryo-EM structure for clathrin in a D6 basket became available[16]. In this structure, presented in the form of carbon alpha(C_α) coordinates of each amino acid in the protein (PDB code 1XI4 [16]), the terminal domains and part of the light chain peptides now resolved, at 1.25 nm resolution. Similar to the method of Tirado Garcia et al.[66, 67], we approximated amino acids by spheres of radius 0.36 nm. This radius was then increased by 0.3 nm to simulate a hydration layer for calculation of R_H. In this case the amino acids were each approximated by a sphere of radius 0.66 nm. The resulting model contained 5100

Tab. 3.1: Experimental Values and Model Calculations of R_H and R_g of Clathrin Triskelion[a][62]

	R_H (nm)	R_g (nm)
light-scattering measurement[b]	16.9 ± 0.3[c]	22.2 ± 1.8
HYDRO calculation using		
a 21 Å resolution structure	14.8[d]	19.3[d]
with the hydration layer	15.3[d]	
ZENO calculation using		
a 12.5 Å resolution structure	15.4 ± 0.3	19.9 ± 0.06
with the hydration layer	15.8 ± 0.3	

[a]The R_H and R_g of clathrin from light scattering is displayed in the first row. Other rows show theoretical predictions based on clathrin structures taken from cryo-EM measurements of D_6 clathrin basket assemblies.

[b]Data taken on Clathrin in 0.1M MES buffer at pH values of 6.0-7.0.

[c]Values corrected for relative viscosity of buffer $\eta_{relative} = 1.05$.

[d]HYDRO calculations do not have error estimates because it involves a deterministic calculation.

spheres, which necessarily overlapped in order to represent the overall dimensions of the triskelion (shown in Fig. 3.10 f'). HYDRO is limited to 2000 spheres and thus could not be used to calculate R_H in this case. Instead, ZENO [55] was used to perform the hydrodynamic calculations. Results of these calculations are shown in rows four and five of Table 3.1. Due to the complexity of this model convergence is slow, and uncertainties (which are shown next to each calculated value) should be noted.

3.3.3 Calculations Using a Segmented Bead Model

In order to estimate the sensitivity of R_g and R_H to changes in clathrin pucker and leg bending, I used a simpler model of a clathrin triskelion (see Fig. 3.5). While our model differs from the protein structures described above, it still captures essential features and uses recent estimates of the dimensions of the clathrin heavy chain and terminal domain. Each leg was given a length of 52 nm, in correspondence with measurements made from transmission electron microscopy (TEM) images of clathrin [17, 9, 10]. In order to model the solution structures for comparison with DLS data, a bead diameter of 3.0 nm, derived from the partial crystal structure for the clathrin proximal leg (PDB code 1B89[22]), was augmented by an additional 0.6 nm to approximate the effect of a single layer of bound water needed when calculating R_H. Thus, sixteen beads (total diameter 3.6 nm) were used, with centers spaced 3.0 nm apart. In contrast, the calculations of R_g were performed for structures modeled with 16 beads of 3.0 nm diameter, as the layer of bound water should be essentially invisible in static light scattering measurements. One sphere of the same diameter

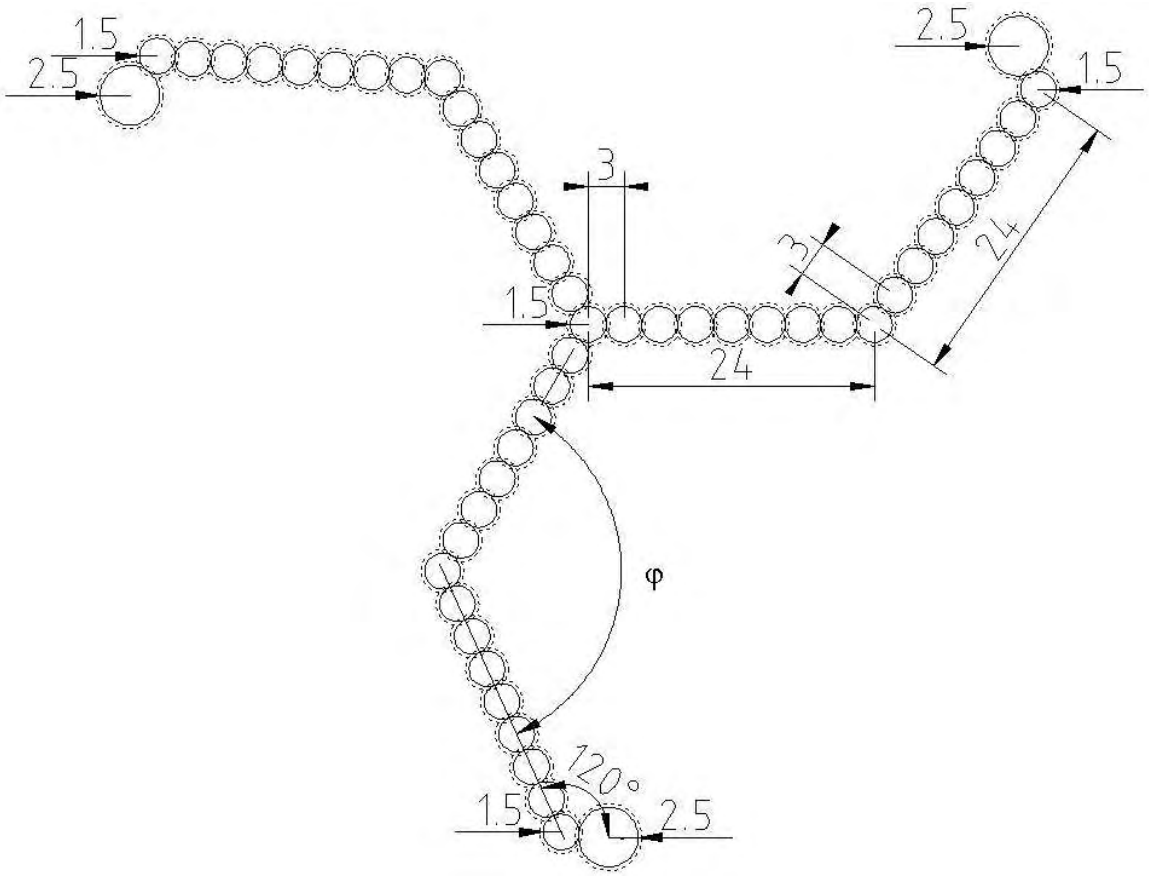

Fig. 3.5: Diagram of simple model of a clathrin triskelion showing dimensions used in our model. Arrows pointing towards a circle indicate the radius of the circle. Lengths are in nanometers and angles are in degrees. The dotted lines indicate the added hydration layer.

was also placed at the hub where the legs join. A larger bead of diameter 5.0 nm was added at the end of each leg to account for the terminal domain, in accord with the aforementioned partial crystal structure (PDB code 1BPO[11]). Again, when calculating R_H, an additional 0.6 nm was added to this bead to account for bound water.

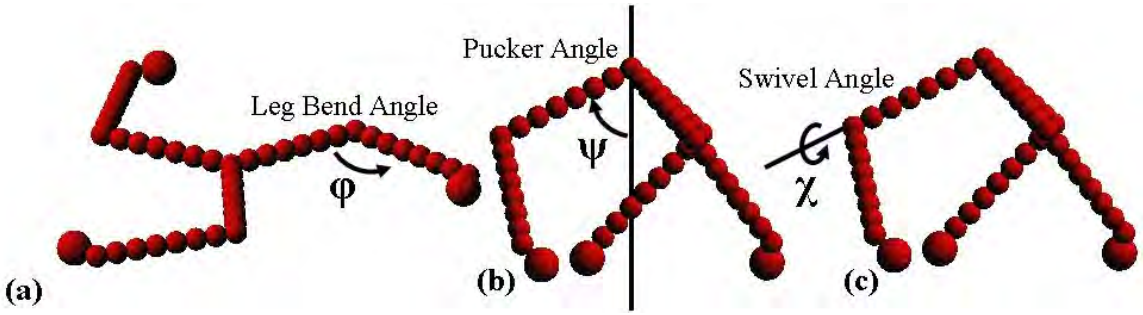

Fig. 3.6: Bead models defining pucker angle , leg bend angle and swivel angles used in modeling. (a)The leg bend angle, φ, and (b) pucker angle, ψ, were varied from 30° to 180° and 30° to 90°, respectively. (c) The swivel angle, χ, was varied parametrically with the pucker angle according to the equation $\chi = 3(\psi - 30)/2$, where ψ is measured from the vertical axis shown in (b).

Each leg included a bend halfway down its length and an out-of-plane pucker measured with respect to a perpendicular line passing through the vertex, defined by the angles φ and ψ respectively. A third angle, the "swivel angle", was varied along the axis of the central leg segment parametrically with the "pucker angle" according to the equation $\chi = 3(\psi - 30)/2$. As the "pucker angle" is varied from 90° to 30° the "swivel angle" varies from 90° to 0° degrees. This equation was selected so that the model structure would change from planar (Fig. 3.6a), similar to that seen in TEM micrographs of individual triskelia on carbon or mica surfaces (Fig. 2.1a), to

puckered (Fig. 3.6b and c), similar to the cryo-EM structures of a triskelion in a clathrin coat(Fig. 2.1b and 2.3) as the pucker angle was varied. Three fold symmetry about the perpendicular was assumed.

By changing the angles φ, ψ and χ, I was able to obtain a quantitative estimate of the sensitivity of R_g and R_H to the conformation of the triskelion as the model changed from planar to puckered. I found that the radius of gyration is much more sensitively dependent on conformation than is the hydrodynamic radius. Both quantities, however, vary considerably for the cases studied, and together provide a distinctive measure of any change in solution conformation. In Fig. 3.7 and 3.8, I show R_H and R_g calculated by HYDRO for different values of bend angle, φ, and pucker angle, ψ and compare them with values measured from experiment (represented by the shaded planes). Note that at $\psi = 90°$ there is no pucker, i.e., the molecule is completely planar, and when $\varphi = 180°$ the legs are completely straight (see Fig. 3.6).

There is considerable degeneracy, in that several pairs of [φ, ψ] are consistent with the measured values. To show this, I combined the information from Fig. 3.7 and 3.8 to produce Fig. 3.9. The upper and lower dark lines (blue and red) indicate the locus of $[\varphi, \psi]$ values commensurate with the measured R_g and R_H respectively. The upper and lower shaded regions in Fig. 3.9 indicate [φ, ψ] pairs yielding values of R_H and R_g, lying within one standard deviation of the respective measured mean values. One can see there is a very narrow band of model conformations (indicated by the overlapping crosshatched region) that are consistent with both the radius of gyration and hydrodynamic radius measured by both static and dynamic light

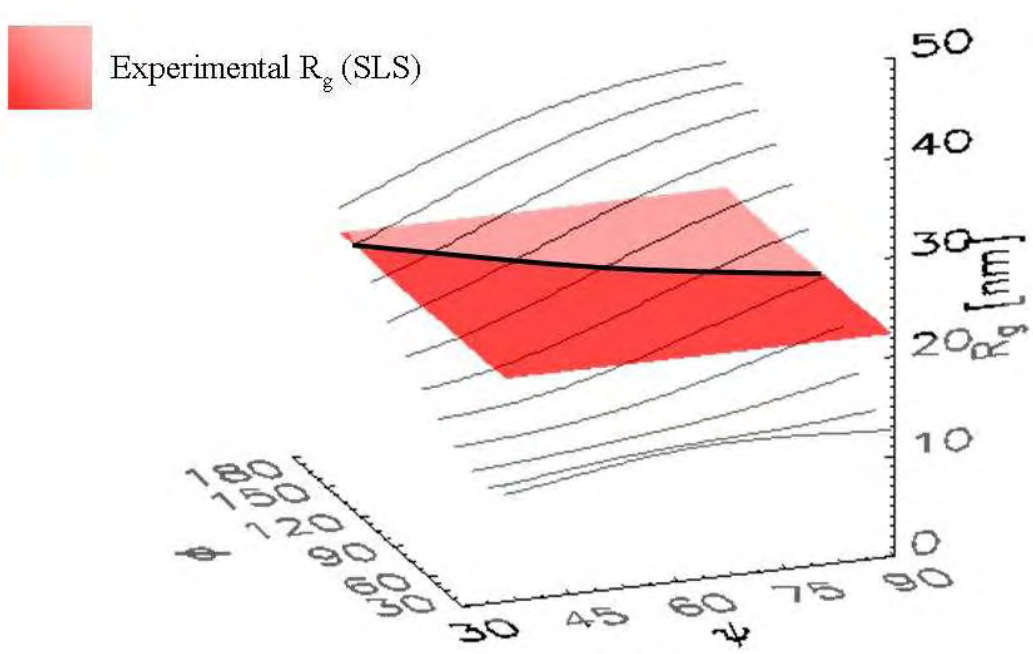

Fig. 3.7: Surface plots of modeled R_g compared with experimental value. Values of R_g calculated for our model are shown in the thin black curves. The shaded surface is a flat plane drawn parallel to the plane defined by the φ, ψ axes (relating to the leg-bend and pucker angles in our model). It shows the average, experimentally-measured, value $R_g = 22.2$ nm. Hence, the dark line indicates the locus of $[\varphi, \psi]$ values commensurate with the measured R_g.[62]

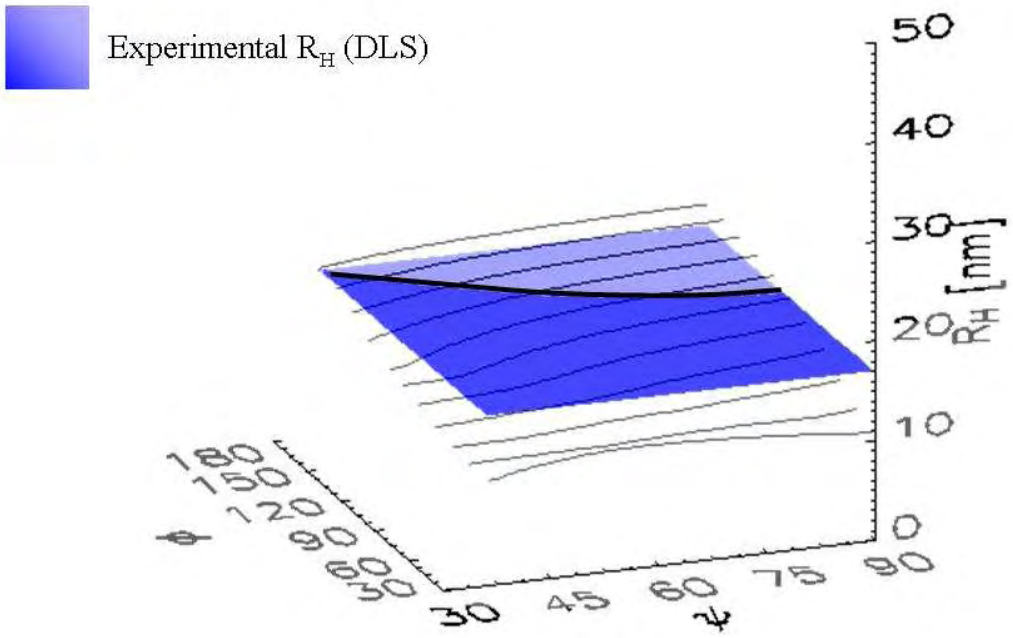

Fig. 3.8: Surface plots of modeled R_H compared with experimental value. Values of R_H calculated for our model are shown in the thin black curves. Similar to what is shown in Fig. 3.7, the shaded surface is a flat plane drawn at the average experimentally-measured value $R_H = 16.9$ nm. The dark line indicates the locus of $[\varphi, \psi]$ values commensurate with the measured R_H[62]

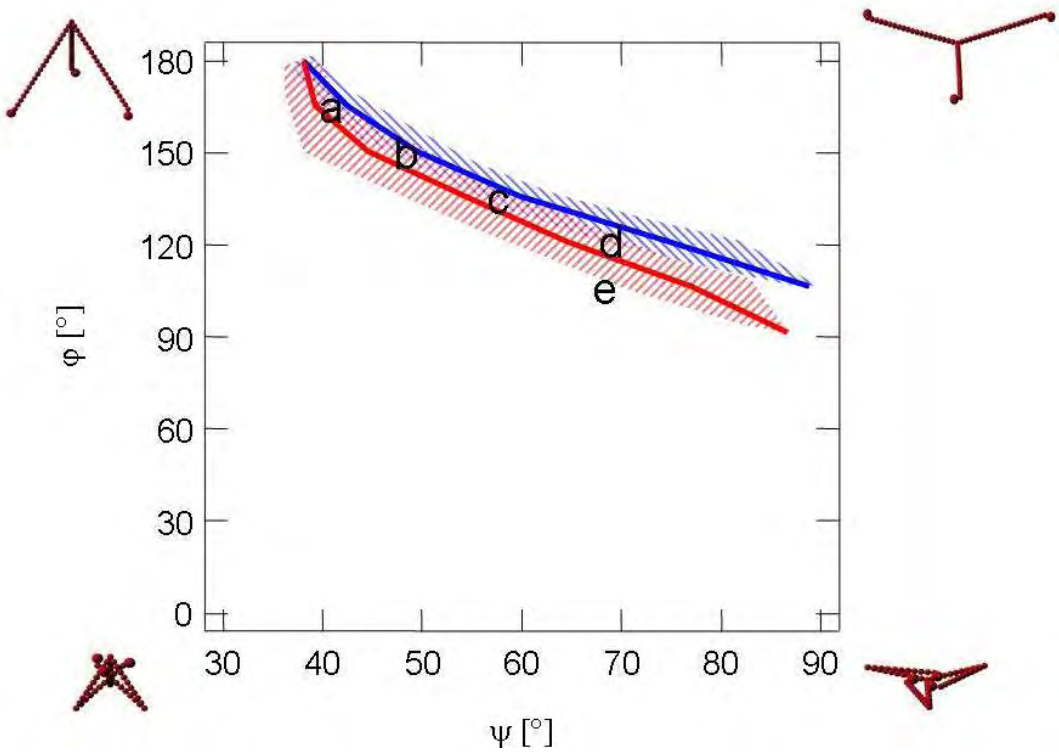

Fig. 3.9: Combination of Fig. 3.7 and 3.8. Values of [φ, ψ] for which calculations based on the segmented bead model match the values for both R_g (lower line) and R_H (upper line) determined from light scattering measurements. The upper and lower dark lines (blue and red) are the respective dark lines of Figures 3.7 and 3.8 indicating the locus of $[\varphi, \psi]$ values commensurate with the measured R_{g} and R_{H}. The upper and lower bands indicate those values of [φ, ψ] lying within one standard deviation of the measured mean values of R_{g} and R_{H}. The region of overlapping bands indicates the range of angles that are compatible with the experimental data from both SLS and DLS. Letters (a-d) indicate models in accord with experimental data, (e) indicates the corresponding [φ, ψ] for a triskelion in a D6 basket (These models are shown in Figure 3.10) ; extreme triskelion shapes corresponding to the [φ, ψ] at the four corners of the graph also are shown. [62]

scattering of free clathrin triskelia in solution. Some compatible structures are shown in Fig. 3.10(a-d). I also indicate the conformations that correspond to the [φ, ψ] values at the corners of the graph. The letter "e" locates the [φ, ψ] pair that was found to be most similar to structures shown in Fig. 3.10e' and 3.10f'.

3.4 Summary and Discussion

Previous TEM investigations indicated a characteristic, pinwheel-like shape for isolated clathrin triskelia absorbed to mica surfaces, the triskelia having a preferred orientation that depends on the buffer in which they are suspended [17, 14]. Based on these studies, quantitative analysis of images of differing handedness [17] led to the inference that free triskelia are puckered, with their hubs raised above the plane defined, by the ends of each of the legs. In those studies, to ensure that they did not associate, the triskelia were suspended in buffers of pH 8.0 when affixed to the mica substrates. I have expanded this work by using dynamic and static light scattering to investigate the conformations of individual clathrin triskelia when free in solution, also extending the investigation to a lower range of pH (pH 6.0 - 6.5). The experimental shape descriptors are the radius of gyration, R_g, and hydrodynamic radius, R_H. By using light scattering, possible structural perturbations of the triskelia due to interactions with EM grids are eliminated.

This study yields three main conclusions. First, I have confirmed that clathrin legs are bent and puckered in solution. To obtain this result, I used HYDRO to calculate R_g and R_H for particular test structures and systematically varied parameters

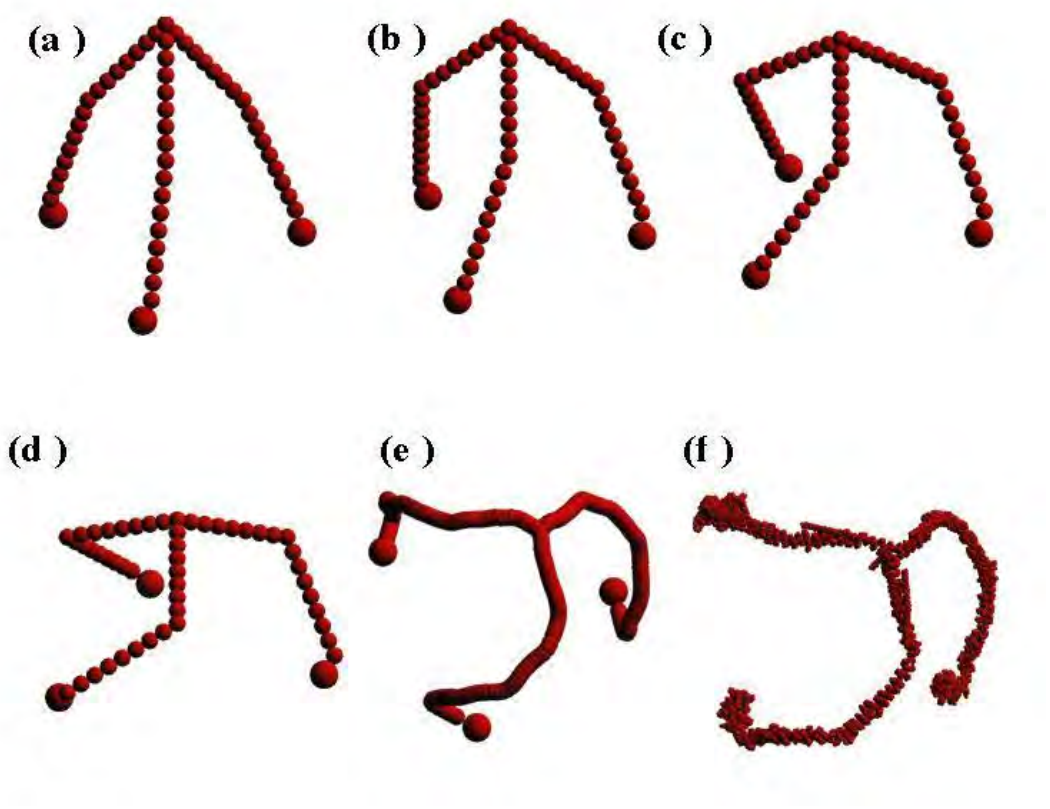

Fig. 3.10: Models that fit experimental measurement using SLS and DLS(a-d) and models based on cryo-EM measurements(e-f). (a-d)indicate the models specified by the corresponding [φ, ψ], and (e-f) show models of a clathrin triskelion derived from 2.1 nm and 1.25 nm resolution cryo-EM tomography of D_6 barrels, respectively[65, 16].

to see which conformations are consistent with the measured data. Yoshimura et al. [68], in a more limited study, investigated triskelia in slightly alkaline solutions (pH 8.0), but modeled them by two dissimilar planar objects when fitting either R_g or R_H. Using this approach, Yoshimura et al. concluded that the solution structures of triskelia in low ionic strength TRIS and TEA buffers (pH 8.0) are not discernibly different, even though triskelia in such samples demonstrate a different handedness when placed onto mica surfaces. In this study I use identical nonplanar models when calculating R_g and R_H to determine the solution conformations of the triskelia.

Second, there is a difference between the solution and basket associated conformations. Cryo-EM of small D6 barrels yields coordinates of a puckered triskelion[65]. Quantitative predictions of R_g and R_H calculated for these shapes differ from values measured for triskelia in solution. However, the structures are relatively close, indicating that the insertion of a triskelion into such an assembly does not require that it undergo a major conformational change. In this regard I note that baskets of differing sizes are found even when they are reconstituted in vitro from purified clathrin alone [14, 69, 13], and that slight differences in leg conformation are observed when one compares triskelia localized to different positions in baskets reconstituted in the presence of AP2[65]. Moreover, the baskets are limited in size (few exceeding ca. 120 nm in diameter), suggesting that, beyond a certain point, a conformation-linked unfavorable energy term may overwhelm favorable interaction energy gained due to the association of the clathrin legs[70, 58]. That these might be mechanical in nature is suggested by the fact that large reconstituted clathrin structures are observed only

when the triskelia lack distal segments, in which case leg bending is suppressed and large, essentially flat, structures are observed[71].

Finally, significant conformational changes in clathrin triskelia are not observed across the pH range known to make the difference between clathrin cage assembly and disassembly. It has been observed that, within experimental error, R_g and R_H do not change as the pH is lowered from 7.0, where basket formation is inhibited, through the "assembly transition" (occurring at pH 6.5 $\rightarrow$ 6.3), to a value (pH 6.0) where essentially all triskelia are rapidly incorporated into baskets. One might have expected conformational changes to occur as conditions were varied from those that disfavor basket assembly to those which induce polymerization. However, as previously mentioned, basket formation is affected by attractive interactions between clathrin legs that are resisted by leg bends and other triskelion shape changes. Thus, our failure to find significant conformational change is important, as it suggests that lowering the pH may affect the mechanical properties of the triskelia (e.g., the rigidity) or enhance inter-leg interactions that favor basket formation (e.g., by changing the ionization state of histidine groups[72]).

The light scattering methods employed in this study are not sufficiently sensitive to demonstrate very small (sub nanometer) changes in the solution structure of the triskelia. Other methods could provide information about triskelia in solution. However, several that immediately come to mind are not yet developed to the point where they can be usefully employed to study molecules of such size. NMR spectroscopy, for example, currently is limited to proteins whose molecular weight does not exceed ca. 40-50 kDa, whereas triskelia have a molecular weight that is an order of magnitude

greater ($>$ 600,000 Da). Similarly, cryo-EM is problematic due to the irregular shapes of the triskelia and the slenderness of their legs. Analytical ultracentrifugation in effect provides the same information as does DLS (the Svedberg coefficient, S, which is obtained by fitting sedimentation data to the Lamb equation[30], can be directly related to R_H) but, by itself, does not yield enough information to discriminate between various model conformations. Measurements of R_g (which is more sensitive to triskelial shape than is R_H) are needed if one wishes to narrow the range of possible structures and thereby eliminate uncertainty about triskelial shape. By using DLS, one is able to determine R_H on the same sample as that used for the measurement of R_g.

In summary, this study utilizes DLS and SLS techniques, in a consistent way, to study the solution structures of clathrin triskelia. It is concluded that a triskelion has an intrinsic pucker, that the solution structure differs from that of a triskelion in a D6 basket, and that large conformational changes do not occur as the pH is lowered below the assembly transition.

The wavelength of light is just short enough to determine information about the triskelion structure. In the next chapter I will describe small angle neutrons scattering (SANS) experiments that were performed in isolated clathrin triskelia in solution. These measurements will be similar to SLS but will extent to higher q probing the triskelion on shorter length scales than light.

4. SMALL ANGLE NEUTRON SCATTERING FROM CLATHRIN IN SOLUTION

4.1 Introduction

In addition to visible light scattering techniques that probe the diffusion coefficient and radius of gyration of a molecule, we can use other methods to noninvasively measure physiochemical properties of clathrin that are related to its structure in solution. In this chapter I describe small angle neutron scattering (SANS) experiments performed on clathrin triskelia. The scattered intensity measured by static light scattering, described in Chapter 3, gave the radius of gyration when fitted by a Gaussian model function: $I(q) \approx e^{-\frac{1}{3}(R_g q)^2}$. SANS allowed us to measure the continuation of the scattering function beyond the Gaussian Guinier regime. Since the neutron wavelength is much smaller than that of light (5.5Å for SANS as opposed to 5145 Å for SLS), it was hoped that measuring the scattering function at higher resolution might give us more information about the protein structure than the R_g, which is the second moment of the particle distribution function, $\rho(\vec{r})$ defined in Section 2.2.7.

In early diffraction studies, Debye and others, using Eq. 2.16, evaluated the scattering function in a few simple cases for molecules of different shapes which were suspended in solution[36, 34, 35]. They found, for example:

1. A uniform and spherical scattering density of radius a, (See Eq. 2.17.):

$$P(q) = \frac{9(\sin(qa) - qa\cos(qa))^2}{q^6 a^6}; \tag{4.1}$$

2. A non-uniform spherical scattering density with a Gaussian radial mass distribution function of width R_g, $\rho(r) = e^{-\frac{r^2}{2R_g^2}}$, corresponding to a random coil (or Gaussian chain):

$$P(q) = \frac{2}{R_g^4 q^4}\left(e^{-R_g^2 q^2} - (1 - R_g^2 q^2)\right); \tag{4.2}$$

3. An infinitely thin rigid rod of length L:

$$P(q) = \frac{2}{qL}\int_0^{qL} dx \frac{sinx}{x} - \frac{1}{(qL/2)^2} sin(qL/2)^2. \tag{4.3}$$

Cases 1 and 3 are useful for some particles that have well-defined spherical and rod-like shapes, such as some viruses. Case 2 is useful for very flexible objects of any type (most studied are different types of linear, star and branched polymers) where dynamics and fluctuations are on very fast timescales compared with measurement or diffusion times.

Others have calculated the effect of flexibility for a semi-flexible linear chain[73] and found a transition from q^{-2} behavior of the scattering function for a random coil (case 2) to the q^{-1} of a rigid rod (case 3). The transition occurs at a value of q that is related to the persistence length (a measure of flexibility) of the linear chain. This connects the limiting cases between case 3 and case 2 with increasing flexibility.

For a number of reasons, the clathrin triskelion is distinct from many other macromolecules and proteins that have been studied by neutron diffraction. The first reason is that the clathrin triskelion is a very large macromolecular complex

containing about 92,000 atoms. Another is that the clathrin triskelion has a three armed pinwheel-like tertiary structure, as mentioned earlier. Because of the triskelion's large size and strange shape, it may be expected that the previous models, originally investigated by Debye and others and frequently used to interpret experimental measurements on smaller molecules of a simpler geometry, may not fit. It was thought that a more appropriate model may be required based on other structural measurements performed with other techniques on triskelia in various environments.

The angular dependence of the scattered intensity from a randomly oriented molecule of arbitrary shape, approximated as a collection of N identical scatterers, can be calculated using the Debye equation (See Section 2.2.5),

$$P(q) = \sum_{i=1}^{N} \sum_{j=1}^{N} f_i(q) f_j(q) \frac{\sin(qr_{ij})}{qr_{ij}} \tag{4.4}$$

where $f(q)$ is the scattering amplitude of a single scatterer. For a sphere of radius a, $f(q)$ is given by Eq. 2.17, $f(q) = 3(\sin(qa) - qa\cos(qa))/(qa)^3$ and Eq. 4.4 becomes Eq. 2.20 of Section 2.2.

Neutron diffraction has been used to verify protein structures obtained by other techniques. For example, one can compare structures measured by X-ray diffraction from crystallized samples or electron microscopy with that of neutron scattering from protein in solution where the molecule is in a more natural environment[74, 75].

4.2 Experimental Methods

Small angle neutron scattering (SANS) measurements were performed on purified clathrin triskelia in solution in order to extend the SLS measurements, discussed

in Chapter 3, into a region of higher q-values, corresponding to shorter length scales. Measurements were made on samples of purified clathrin in disassociation buffer, 0.5 M Tris (2-amino-2-hydroxymethyl-1,3-propanediol) pH 7.0, in H_2O and in basket assembly buffer, 0.1 M MES (2-(4-morpholino)ethanesulphonic acid), at low and high pH (6.0 and 7.0) in H_2O. This experiment was done in order to determine if the scattering function from purified clathrin triskelia in solution was dependent on buffer type or pH. In order to reduce incoherent scattering, SANS measurements were made on clathrin in D_2O buffers, making it possible to take measurements at even higher q-values than in an H_2O solvent.

4.2.1 Sample Preparation

Clathrin protein was purified from coated vesicles of bovine brain as discussed in Chapter 3.

Clathrin samples in D_2O were dialyzed into D_2O buffers at a pD^* value of 7.0 for the sample measured in disassociation buffer, and at pD^* values of 6.0 and 7.4 in basket assembly buffers. It should be noted here that pD^* values reported were all measured by a glass electrode pH meter and so read at a value of 0.4 units below the actual $pD(= -\log_{10}|[D^+]|)$:

$$pD = pD^* + 0.4\ units.$$

On the day of the SANS experiment, all samples were centrifuged for 15 min at 83,000 rpm using a Beckman MLA130 rotor (corresponding to an acceleration of $\sim 400,000\ g$). The concentration after centrifugation was between 2.6 and 3.1 g/L.

Samples were then diluted into the buffer used for dialysis to a protein concentration of between 1.5 and 2.4 g/L. This was done so that the volume would be $\sim$ 300 μL, filling up the entire cuvette.

For low and high pH MES samples, purified clathrin samples at a concentration of 3.2 g/L in 0.5 M Tris buffer at pH 7.0 + 3 mM DTT were dialyzed overnight, at a temperature of $4°C$ in 10,000 molecular weight dialysis cassettes (Pierce Bio), into 30 mL of 0.01 M Tris buffer at pH/pD* 8.0 + 3 mM DTT in H_2O/D_2O. Immediately prior to the experiment, 1/10 volume of 1M MES buffer was added to a clathrin sample. For samples in D_2O, 1/10 volume of 1M MES in D_2O was added to the clathrin sample. Each sample was then loaded into a 1 mm thick quartz cuvette.

For the clathrin sample in 0.5 M Tris D_2O buffer, a purified clathrin sample at a concentration of 3.2 g/L in 0.5 M Tris buffer of H_2O at pH 7.0 + 3 mM DTT was dialyzed overnight, at a temperature of $4°C$ in 10,000 molecular weight dialysis cassettes (Pierce Bio), into 30 mL of 0.5 M Trizma in D_2O. The sample was then loaded into a 1 mm thick quartz cuvette.

0.5 M Tris D_2O buffers were made by the addition of Trizma (Tris-HCl © Sigma-Aldrich) to 99.9% D_2O. This dissolved to a pH of 5.6 and allowed us to add NaOH dissolved in D_2O to get to pD* 7. The final buffer contained $\sim$ 97% D_2O due to the inadvertent addition of 3 mL of H_2O and 750 μL of 0.5 M DTT in H_2O to 120 mL of D_2O buffer (experimental error).

Concentrations were measured by UV absorbance at a wavelength of 280 nm as described in Chapter 3.

4.2.2 Light Scattering

All SLS and DLS measurements of clathrin in solution were made as described in Chapter 3.

4.2.3 Neutron Scattering

Experiments were performed on the NCNR NG-3 30 meter Small Angle Neutron Scattering Instrument. A diagram of the instrument is shown in Fig. 4.1. The wavelength of neutrons used was 5.5 ± 0.4Å (0.8Å FWHM) at sample to detector distances of both 13.16 and 5.00 meters. q-values ranged from ≈ 0.003 to 0.04 Å^{-1} and 0.01 to 0.1 Å^{-1}, respectively. Multiple data files were recorded for acquisition times between 15 minutes to 1 hour and were later added together. Total data acquisition times for each sample were between 3 and 5 hours per sample. Protein concentrations were between 1.5 and 2.4 g/L. Neutron transmission was 83-86% in D_2O and $\sim 52\%$ in H_2O. Typical raw data are shown in Fig. 4.2 for a detector distance of 13.2 m and in Fig. 4.3 for a detector distance of 5m. Circular averages were done in order to improve data statistics. Background scattering was measured for empty cuvettes (for 1 hour) and buffers in both H_2O (4 hours) and D_2O (2 hours).

4.3 Results

4.3.1 Testing D_2O as a Solvent for Clathrin

Due to the absence of hydrogen in the sample, higher signal to background can be achieved during a neutron scattering experiment when D_2O (^2H_2O) is used as a

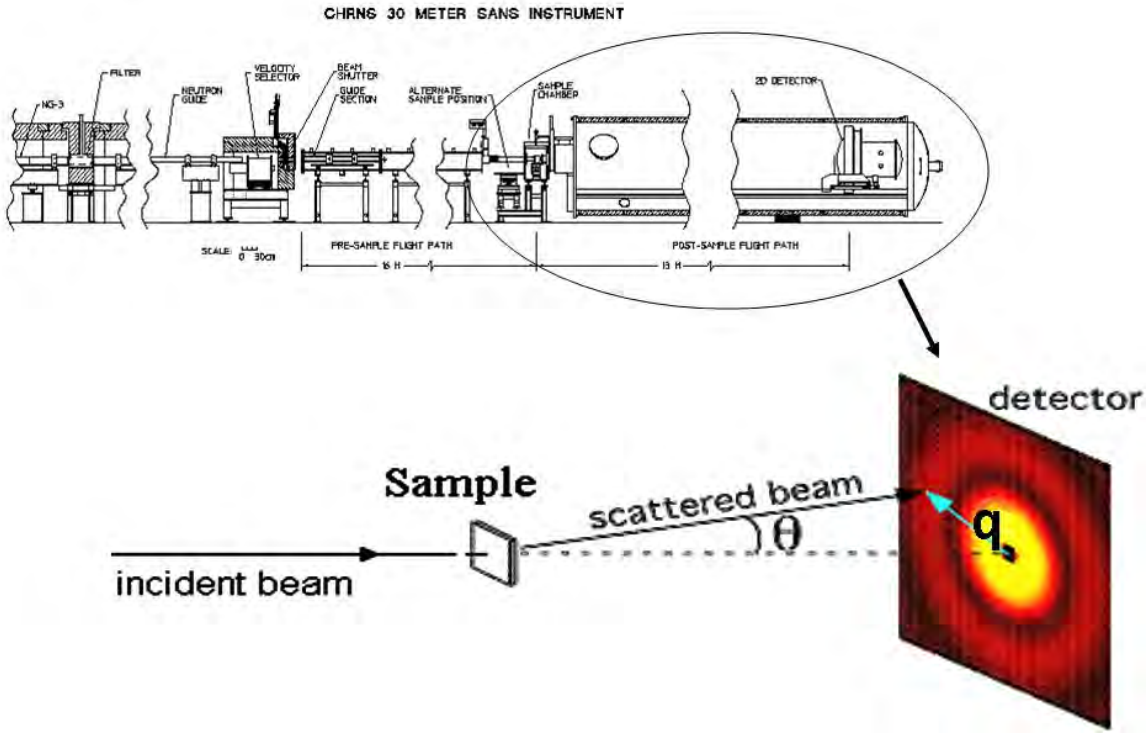

Fig. 4.1: (top)Diagram of SANS NG3 instrument used: "Cold neutrons from the source pass through a multi-disk mechanical velocity selector with variable speed and pitch, enabling both the mean wavelength and wave-length resolution to be varied. The monochromatic beam is collimated by circular pinhole irises in a 15 m long evacuated pre-sample flight path. The source iris can be preceded by up to eight 1.5 m guide sections that can be shifted in or out of the beam. The beam divergence and flux on the sample can thus be varied by changing the effective source-to-sample distance from 4 m to 16 m in 1.5 m increments. The Instrument is designed to cover a wide q-range, from 0.0015 to nearly 0.6 $Å^{-1}$." Figure and caption taken from http://www.ncnr.nist.gov/instruments/ng3sans/index.html. (bottom) A schematic diagram of the physically relevant details of the SANS instrument.

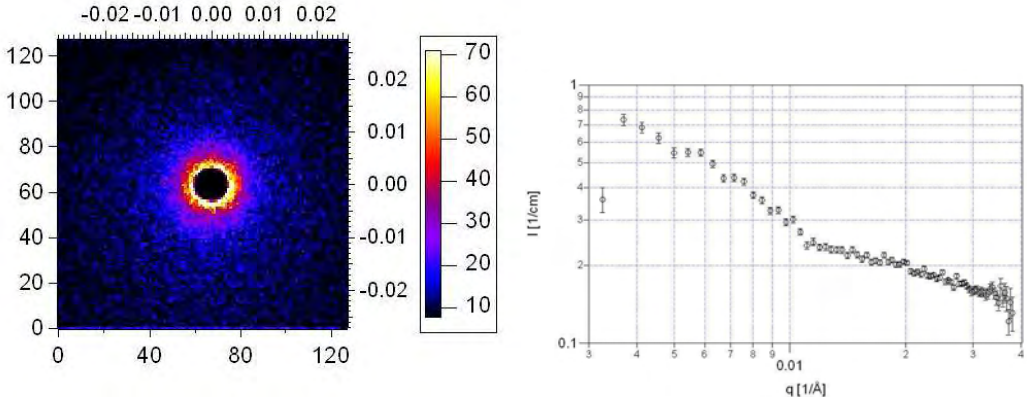

Fig. 4.2: Raw SANS data of clathrin triskelia in D_2O at a detector distance of 13 meters. A 1 hour exposure of clathrin triskelia susspended in 0.5 M Tris pH 7.0 D_2O buffer. The color scale is in counts/second. The bottom and left axes in the image are pixel numbers. The top and right axes are in units of $Å^{-1}$. The sample to detector distance was 13.16 meters. The graph on the right shows a circular average of the image on the left in absolute units (1/cm). The protein concentrations was measured to be 1.45 mg/mL by UV absorbance.

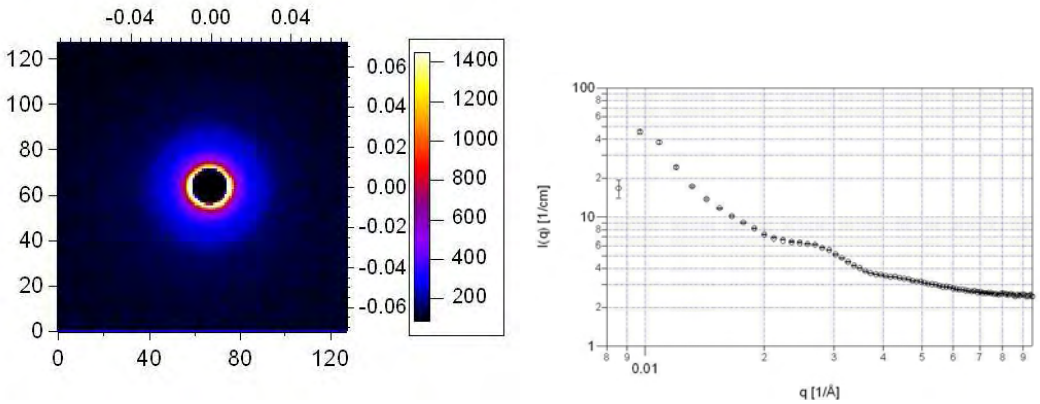

Fig. 4.3: Raw SANS data of clathrin baskets in D_2O at a detector distance of 5 meters. A 1 hour exposure of clathrin baskets in 0.1 M MES pH 6.0 D_2O buffer. (Compare to Fig. 4.2) The sample to detector distance was 5 meters. The protein concentration was measured to be 2.28 mg/mL by UV absorbance.

solvent. This is because D_2O has a smaller incoherent scattering cross section than H_2O for neutron scattering. (See Table 2.1) The result is that a lower background scattering is achieved in the angular intensity profile. This can be readily seen by the higher transmission efficiency in the D_2O samples, which was ∼83-86% in D_2O and $\sim 52\%$ in H_2O.

D_2O has many of the same properties as H_2O, but it has been found to affect some biological macromolecules[76, 74] and therefore should be tested as a solvent in each particular case. It was necessary to check that clathrin remains folded properly and maintains its functionality when D_2O is used as a solvent. For this reason light scattering experiments were performed on clathrin triskelia dissolved in D_2O buffers.

For comparison, neutron scattering was performed on protein samples in H_2O as well. In those samples our sensitivity was limited to q values for which the neutron

scattering intensity due to the presence of protein was on the order of or larger than the incoherent scattering due to the solvent. In practice, this limited us to q values below 0.025 $Å^{-1}$ as can be seen by the magnitude of the error bars on the experimental SANS data shown in Fig. 4.6.

Clathrin Maintains Its Structure in D_2O

To test the stability of clathrin in D_2O, both static and dynamic light scattering measurements were performed. In light scattering, different solvents can affect measurements through changes in relative index of refraction between the solvent and solute, which will affect q dependence as well as the intensity of the scattered light. Another is a change in viscosity, which will affect the diffusion of the particles and affect the DLS measurement. The index of refraction for D_2O, shown in Table 2.1, is very close to that of H_2O. The viscosity was found to increase by 22% in D_2O, by dynamic light scattering from polystyrene spheres of radius $\sim 60nm$, consistent with values found in literature[77, 78]. Light scattering measurements of clathrin triskelia in both solvents were comparable once a correction was made in the DLS measurement for a difference in viscosity between the two solvents. Viscosity-corrected results of the light scattering measurements made on clathrin in both H_2O and D_2O buffers are shown in Table 4.1.

Clathrin Assembles Baskets in D_2O

In H_2O (and in 0.1 M MES buffer), clathrin triskelia will rapidly assemble below a pH of 6.5[14]. In order to test whether clathrin will function in a D_2O solvent,

Tab. 4.1: Experimental Values of R_H and R_g of Clathrin in 0.5 M Tris in both H_2O and D_2O solvents.[79]

	R_H (nm)	R_g (nm)
H_2O	15.7 ± 0.1^b	22.0 ± 0.5
D_2O	15.6 ± 0.2^b	19.5 ± 0.9

[a]Measurements were made by light scattering experiments.

[b]Values corrected for relative viscosity of

D_2O ($\eta_{D_2O} = 1.22$) [77, 78] and buffer, ($\eta_{0.5MTris} = 1.18$).

clathrin basket assembly in 0.1 M MES buffers made with both H_2O and D_2O was monitored by light scattering, at pH and pD* values between 6.0 and 7.2. The extent of basket assembly was monitored dynamically by DLS and by a final pelleting assay in a Beckman ultracentrifuge[80] 24 hours after the initial addition of MES buffer (the initiation of basket assembly).

In order to capture the time dependence of the hydrodynamic radius of clathrin by DLS during assembly, it was necessary to reduce the correlation time from 2 minutes to 5 seconds. If the acquisition time was larger than this, the hydrodynamic radius was overestimated during periods of rapid basket assembly. This caused larger errors in estimation of the hydrodynamic radius, but allowed a rough characterization of size and fraction of polymerization. Both hydrodynamic radius and absolute intensity were monitored. The combination of the scattered light intensity along with the intensity correlation function showed definitively that clathrin triskelia maintained

Tab. 4.2: Fraction of Clathrin in Pellet after 24 hours in H_2O and D_2O.[79][a]

pH/pD*	H_2O	D_2O
7.0/7.2	(0)	(0)
6.5	(0)	0.20
6.2	0.53	0.65
6.0	0.60	0.89

[a] Measurements of protein concentration were made by UV absorbance at a wavelength of 280 nm. The absorption coefficient of ϵ_{280} was estimated to be 1.07 $g^{-}1\ cm^{-1}$ from the amino acid sequence[61].

both structure and function in D_2O. Scattered intensity and hydrodynamic radius vs. polymerization time are shown in Fig. 4.4.

After 24 hours each sample was centrifuged at 10,000 g for 15 minutes to estimate the fraction of protein assembled into clathrin baskets. Polymerized fraction vs. pH is shown in Table 4.2 and Fig. 4.5[80].

Clathrin baskets assembled faster in D_2O buffer than in H_2O buffer for a given measured pD* value and pH. Clathrin could also assemble at higher pH values in D_2O than in H_2O buffers (e.g. pH/D* 6.5). This is even more surprising, given the fact that the glass electrode meter measures a lower pD* value than the actual pD, as mentioned in Section 4.2(e.g. pD*=6.5 $\rightarrow$ pD=6.9). This promotion of basket assembly by D_2O is still not fully understood and is a topic of active research.

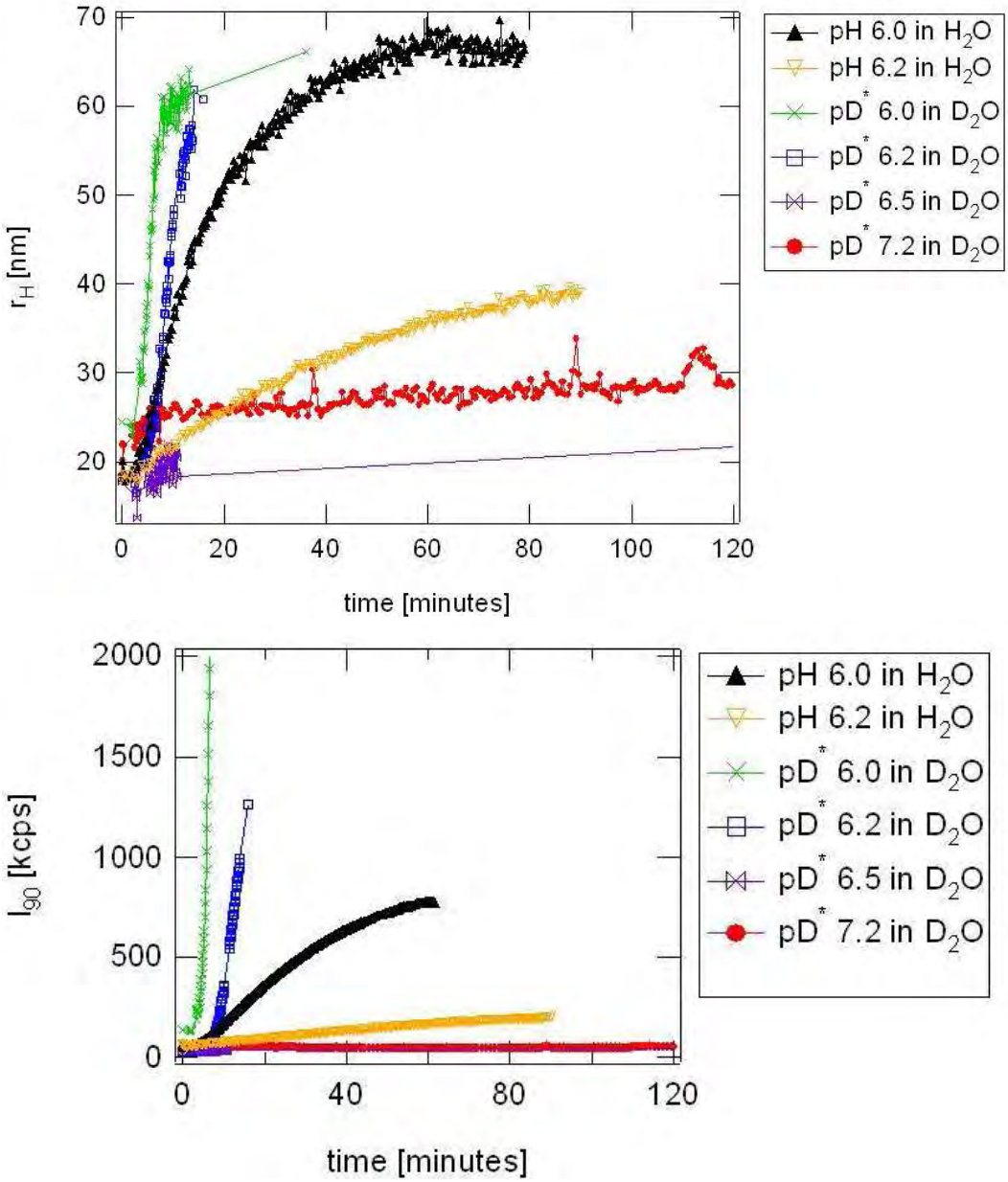

Fig. 4.4: Light scattering measurements show clathrin assembles faster in D_2O buffer than H_2O buffer at pH and pD* values of 6.0 and 6.2. (as measured by meter). Baskets do not assemble in D_2O or H_2O at pH or pD* values of 6.5 and 7.2. For comparison, clathrin basket assembly is shown in H_2O at a pH value of 6.0 and 6.2. All experiments were conducted at room temperature.[79]

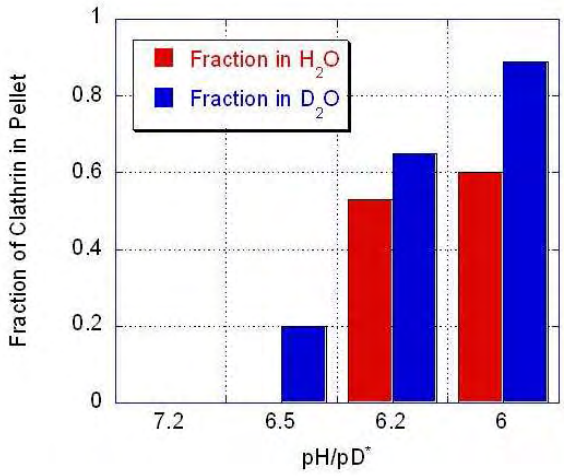

Fig. 4.5: Fraction of clathrin pelleted after 24 hours vs. pH/pD* in H_2O/D_2O (from Table 4.2).[80]

4.3.2 Neutron Scattering and Rigid Models

SANS data were compared with theoretical curves based on rigid bead models of clathrin triskelia identical to those used to fit experimental light scattering data from Chapter 3. Calculations of the scattering function, $P(q)$, were made using Eq. 4.4 for bead models based on the 21 Å and 12.5 Å resolution cryo-EM structures of a clathrin triskelion when assembled into a D6 basket with AP2[65, 16]. As in Chapter 3, for the highest resolution model (PDB Code 1XI4[16]), a bead of radius 3.6 Å was used to represent each amino acid residue resolved in the structure. Each bead is centered at the coordinates of the C_α atom of the corresponding amino acid.[67, 66]

A lower resolution model used consisted of 52 beads. This model was used to estimate the affect of changes in conformation of the triskelion (see Fig. 3.6) on the calculated theoretical scattering function. The 16 beads in the leg of the low

resolution models have a radius of 15 Å, corresponding to about 12 kDa of protein or approximately 110 amino acid residues. The terminal domains are modeled as slightly larger beads, having a radius of 25 Å each. (See Fig. 3.5.) A visualization of the models as well as results of the calculations can be seen in Fig. 4.6 and 4.7, along with SANS data taken on clathrin triskelia in 0.5 M Tris at pH 7.0 in H_2O.

Two characteristic features are observed after applying Eq. 4.4 to these rigid bead models. They are a minimum and secondary maximum at q values just outside of the central Guinier regime. This dip occurs because of destructive interference of the scattered wave from scatterers in different legs of the triskelion. It does not occur, for example, in calculations done on a single leg (Shown in Fig. 4.8), and seems to be sensitively dependent upon the relative distance and positions between the legs and the out of plane pucker of the triskelion. (See Fig. 4.9) This feature seems to decrease in prominence as the molecular model is flattened parametrically. A noteworthy observation from the SANS data in Fig. 4.6 and 4.7 is that this dip, calculated for the puckered triskelial models, does not exist as prominently in the experimental data. It occurs only as a shoulder or plateau just outside the low q, "Guinier regime." This leads us to speculate on the reasons for the discrepancy between experiment and our models.

Another interesting feature of the experimental SANS data is the power law decay of the scattered intensity noted for large values of the magnitude of the scattering vector q. It falls off somewhere between the $1/q$ behavior of a rigid rod model and the $1/q^2$ behavior of a random coil ($\sim q^{-1.23}$ in D_2O). This can be seen in Fig. 4.6 where the solid black lines indicate the slopes.

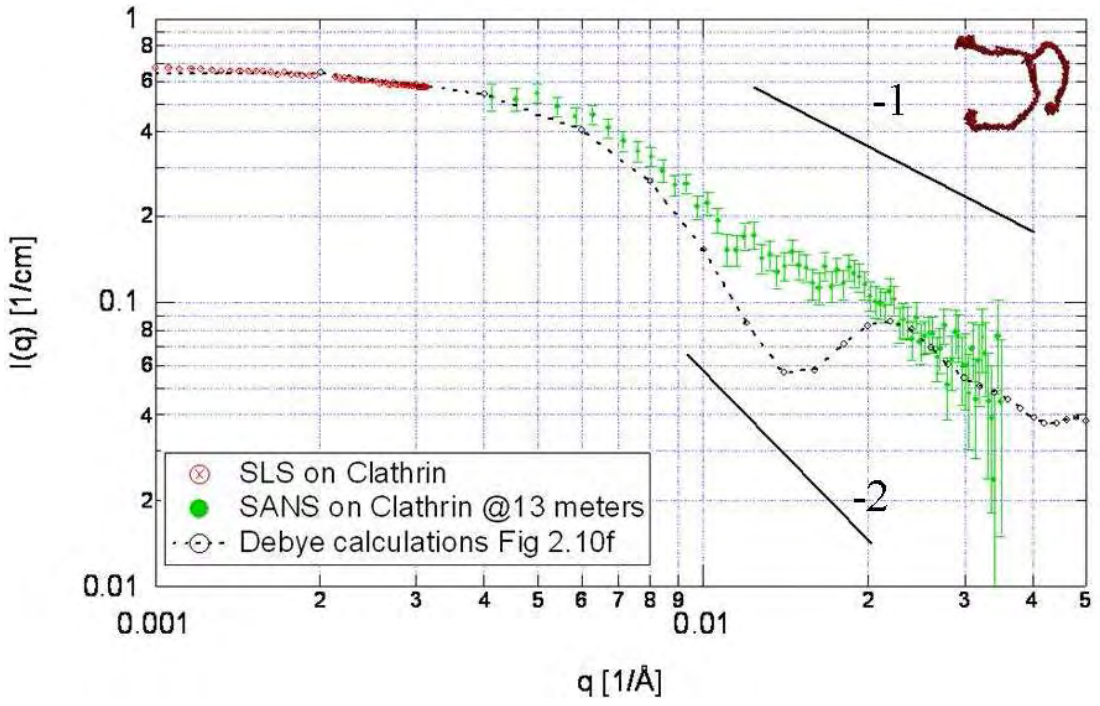

Fig. 4.6: SLS and SANS data and calculated scattering function of a cryo-EM model. SLS and SANS data of clathrin in 0.5 M Tris in H_2O. A constant background of H_2O buffer was subtracted of $\sim$ 1 cm^{-1}. The SLS data (in counts per second) was scaled to the SANS data. The dotted line shows a scattering function calculated using Eq.4.4 for cryo-EM structure of clathrin in a D6 basket. The model is shown in the upper left-hand corner. The inset shows the graph in a linear plot. The calculated scattering function, which is normalized to 1, was multiplied by a factor of 0.63 cm^{-1} in order to get best agreement with both SLS and SANS data. The solid black lines indicate the power laws $1/q^2$ and $1/q$ as indicated corresponding to the large q scaling of a flexible Gaussian chain model and a rigid rod respectively.

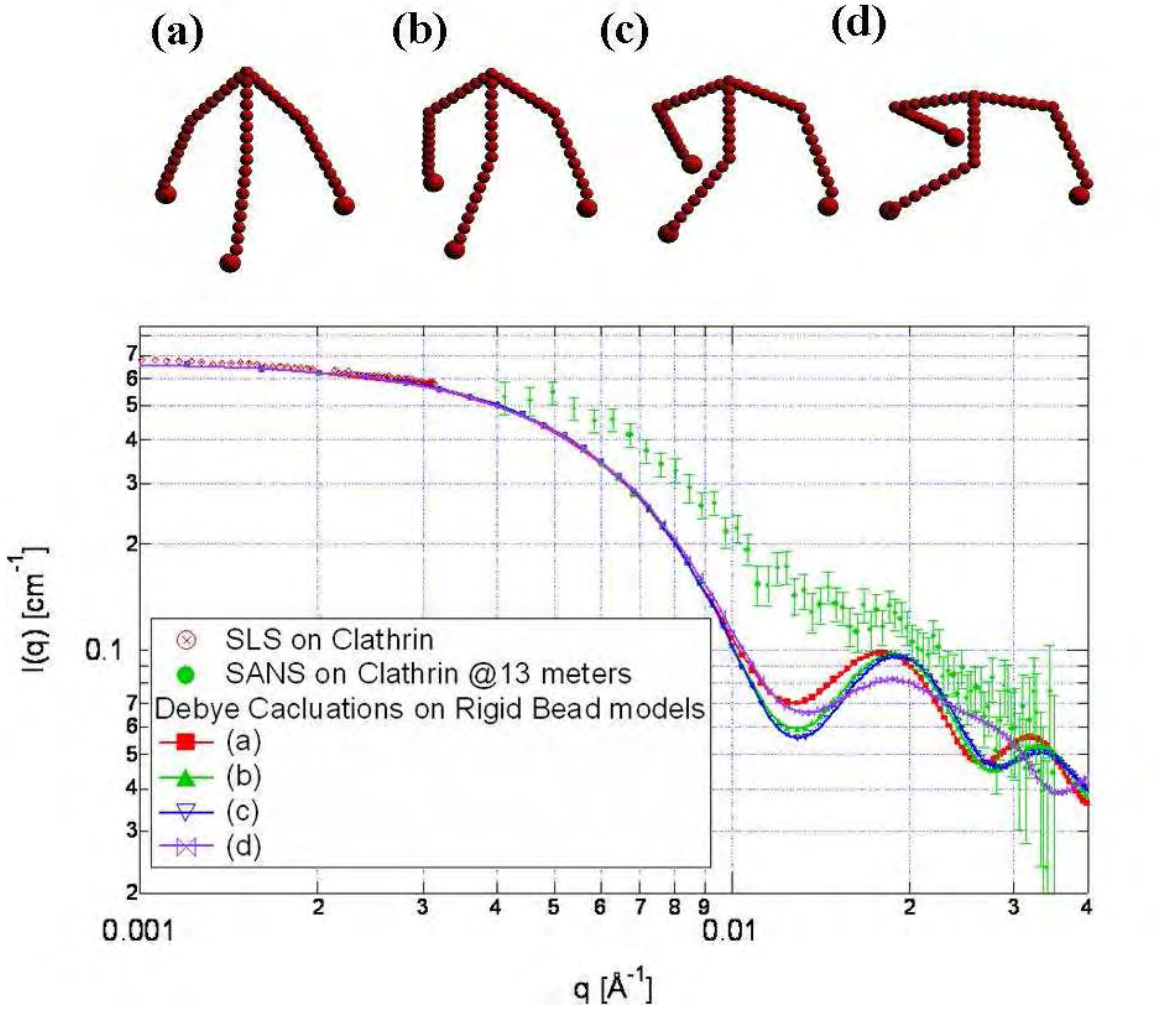

Fig. 4.7: SLS and SANS data and calculated scattering function of rigid bead models of clathrin. (Consult Fig. 4.6.) The dotted line shows a scattering function calculated using Eq. 4.4 for 4 models of a clathrin triskelion made up of 52 beads and shown above in (a)-(d). These are the models that fit the DLS and SLS measurements from Chapter 3. The inset shows the graph in a linear plot. The calculated scattering function which is normalized to 1 was multiplied by a factor of 0.66 cm^{-1} in order to get best agreement with both SLS and SANS data.

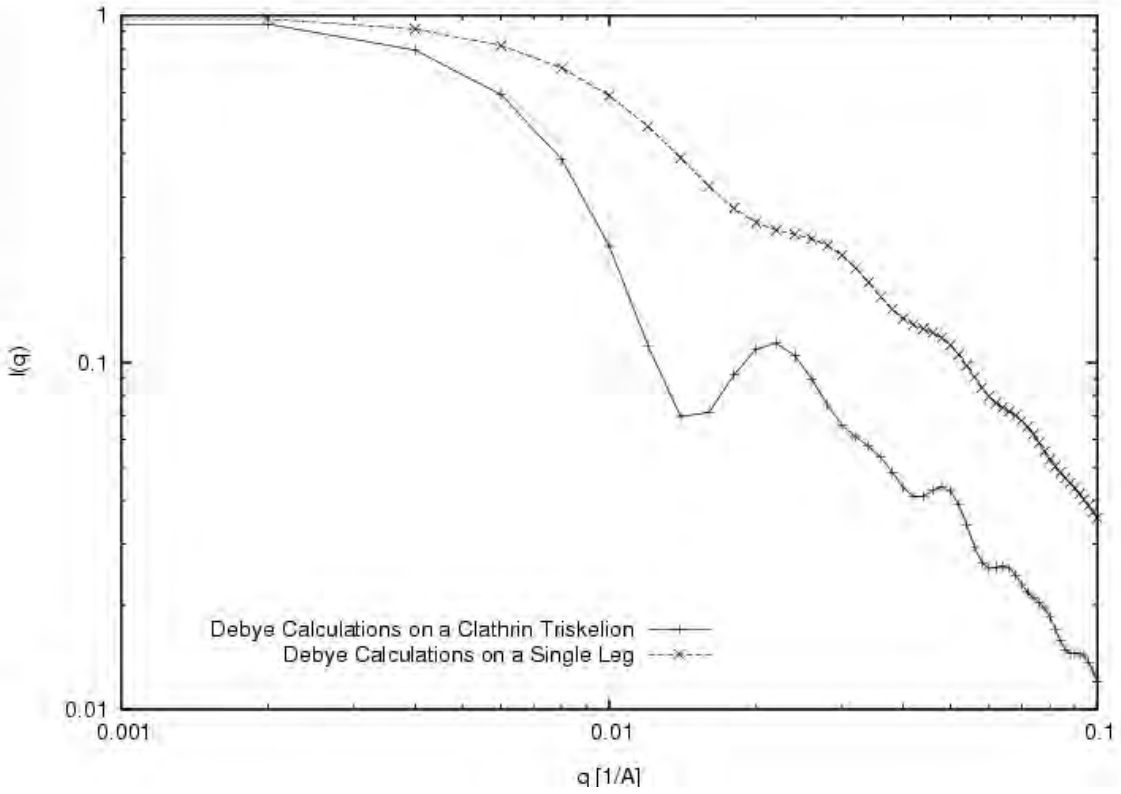

Fig. 4.8: The calculated scattering function of triskelion leg (upper curve) and a complete triskelion with three tegs (lower curve). The model used was the a high resolution cryo-EM structure of a triskelion in a D6 basket. The model is shown in Fig. 3.10f and Fig. 4.6.[16]

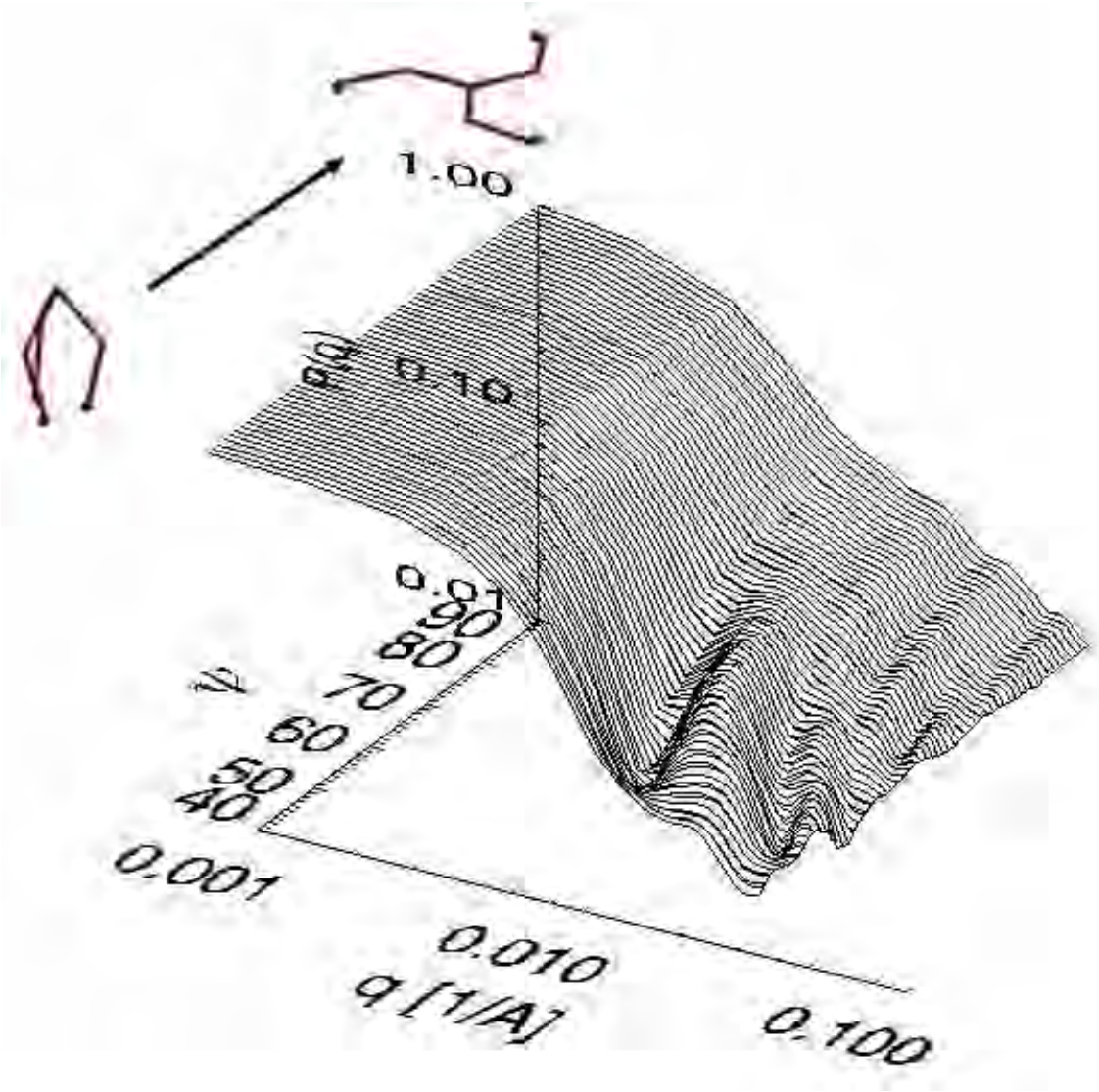

Fig. 4.9: Surface plot variation of scattering function with pucker angle. A surface plot showing changes in the calculated scattering function for a 52 bead triskelion model as the pucker angle, ψ, is varied from 90° to 30° and swivel angle, χ , is varied parametrically with the pucker angle according the the equation $\chi = 3(\psi - 30)/2$. The leg bend angle is fixed at $\phi = 120°$.

4.3.3 Flexible Models

For a completely flexible model of a linear chain polymer, the scattering function is given by Eq. 4.3. A flexible model of a triskelion as a fully flexible molecule of three connected linear chains takes into account interference between scatterers in different chains as well as the same chain. The scattering function still falls of asymptotically as q^{-2} at large q and is described by the following equation (from Teraoka[81]):

$$P_3(q) = \frac{6}{R_g^4 q^4}\left(e^{-\frac{1}{3}R_g^2q^2}(e^{-\frac{1}{3}R_g^2q^2} - 1) + \frac{1}{3}R_g^2q^2\right) \tag{4.5}$$

This flexible model for a clathrin triskelion does not fit SANS data at large values of q. This can be seen in fits to SANS data from clathrin suspended in H_2O buffer, Fig. 4.10 (0.5 M Tris pH 7) and more clearly in fits to SANS data from clathrin suspended in D_2O buffer, Fig. 4.11 (0.5 M Tris pH 7) where data could be taken at distances of 13 meters and 5 meters, measuring the scattering function out to $q = 0.1\mathring{A}^{-1}$

4.3.4 Smearing Caused by a Wavelength Distribution

In the NG3-SANS instrument, neutrons from the cold source pass through a multidisk mechanical velocity selector with variable speed and pitch, enabling both the mean wavelength and wavelength resolution to be varied since the neutrons wavelength is determined by its velocity by the de Broglie relation: $\lambda = \frac{h}{mv}$. In our experiments the mean wavelength of neutrons was 5.5 Å and the full width at half maximum was 0.8 Å.

In the calculation of model scattering functions already presented, it is assumed that the incident and scattered wave vectors of interfering plane waves, $\vec{k}_0$ and $\vec{k}_s$, have

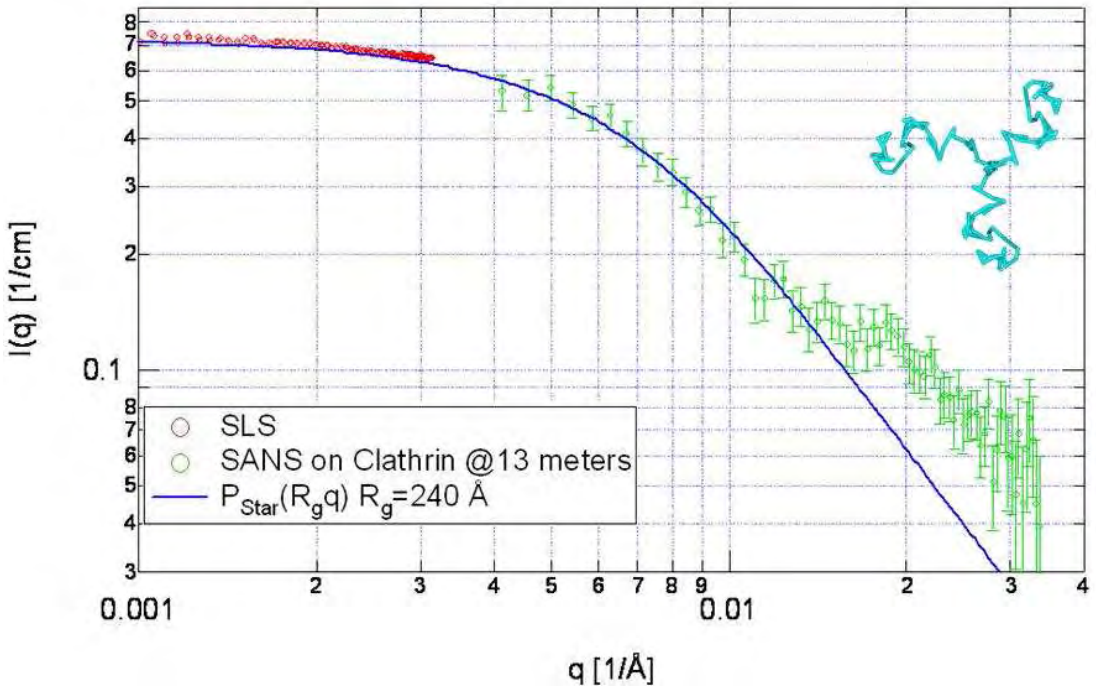

Fig. 4.10: SLS and SANS data in H_2O buffer and fit to flexible three-armed star polymer model. (Compare to Fig. 4.6.) The solid line shows a fit to a calculated scattering function for a fully-flexible three-armed star polymer (Eq. 4.5). Note the deviation of the model function at high values of q.

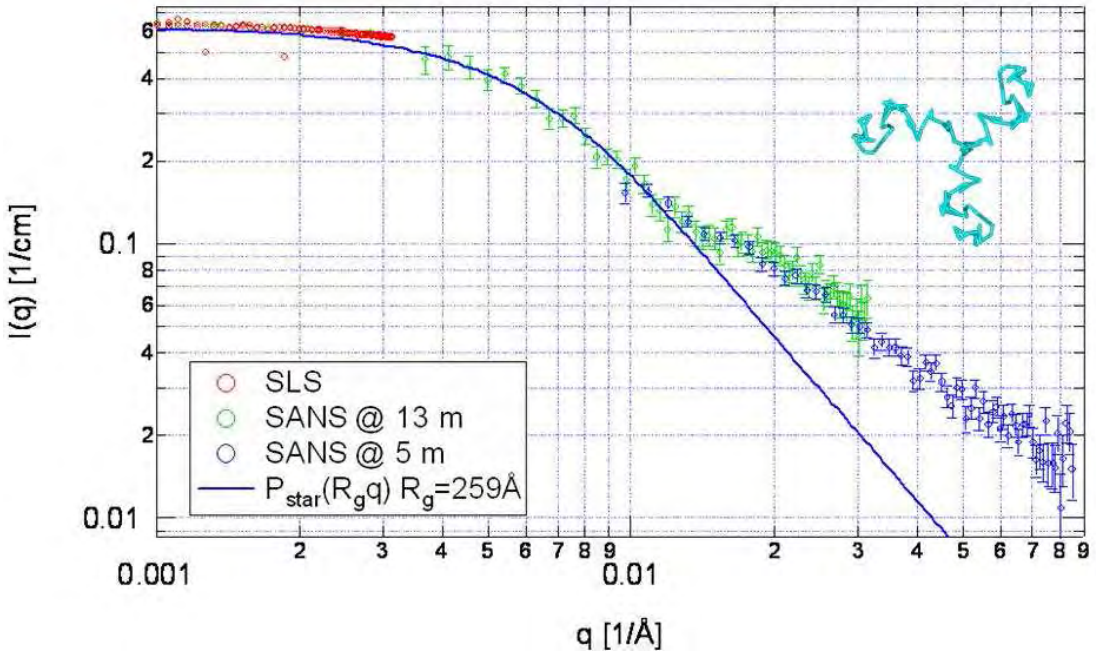

Fig. 4.11: SLS and SANS data of clathrin in D_2O buffer and flexible three-armed star polymer model. SLS and SANS data of clathrin was taken in 0.5 M Tris D_2O buffer. A constant background from D_2O buffer was subtracted of ~ 0.1 cm^{-1}. The SLS data (in counts per second) were scaled to the SANS data. The solid line shows a fit to a calculated scattering function for a fully-flexible three-armed star polymer (Eq. 4.5). Note the deviation of the model function at high values of q.

a fixed magnitude. If instead the incident beam has a wavelength dispersion, then the magnitude of the scattering vector ($q = 2|\vec{k}_0| \sin\frac{\theta}{2} = \frac{4\pi}{\lambda}\sin\frac{\theta}{2}$) will be described by a Gaussian distribution function,

$$\rho(q) = \frac{1}{\sqrt{2\pi\sigma_q^2}} e^{-\frac{(q-q_0)^2}{2\sigma_q^2}},$$

where q_0 is the average wavevector and σ_q is width, and the actual scattering function will be averaged over the diffraction pattern from many wavelengths according to the following equation[31]:

$$P_\rho(q|q_0,\sigma_q) = \int_0^\infty \rho(q')P(q')dq' = \frac{1}{\sqrt{2\pi\sigma_q^2}} \int_0^\infty P(q') e^{-\frac{(q-q_0)^2}{2\sigma_q^2}} dq'. \qquad (4.6)$$

This transformation was applied to the triskelion model function[1] for $q_0 = 5.5$ Å and $\sigma_q = 0.8$ Å . The results may be seen in Fig. 4.12. One can see from this calculation that smearing due to wavelength dispersion is insufficient to explain disagreement between experiment and theory, in the case of rigid bead models shown in Fig. 4.6 and 4.7.

4.4 Summary and Discussion

The SANS experiments described in this chapter are consistent with all of the findings of the light scattering experiments of Chapter 3. We can rule out a rigid planar structure for the triskelion because it is not consistent with light scattering. The radius of gyration of such a structure is too large and does not fit SANS measurements

[1] Here I would like to thank and acknowledge S. Krueger, who kindly performed this calculation for me and produced this graph.

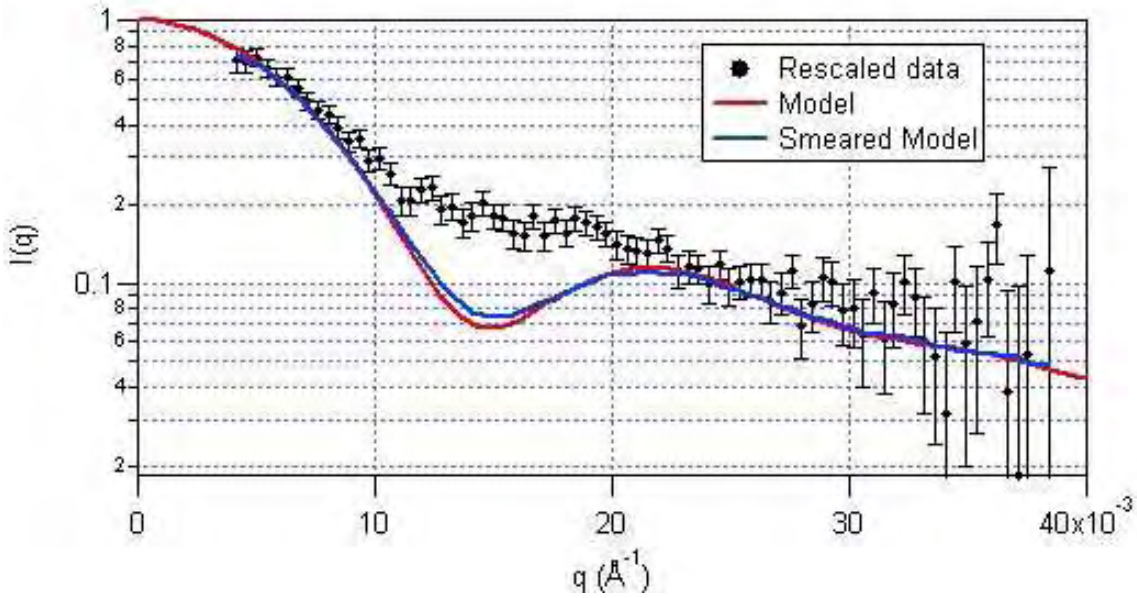

Fig. 4.12: SANS data and smeared scattering function of a cryo-EM model. SANS data of clathrin in 0.5 M Tris in H_2O. A constant background of H_2O buffer was subtracted of $\sim$ 1 cm^{-1}. The red line shows a scattering function calculated using Eq. 2.20 for the cryo-EM structure of Clathrin in a D6 basket. The blue line shows the scattering function smeared using Eq. 4.6. The model and SANS data were normalized to 1.

at small values of q, even though the model calculations might fit at intermediate and large q. Neither will a fully flexible Gaussian chain model adequately describe SANS measurements at large q.

The most consistent models presented in this chapter are the rigid puckered bead models (both from the cryo-EM model and the 52 bead models) which fit both DLS and SLS measurements. These models are also in good agreement with SANS measurements, except for the discrepancy at intermediate values of q. Inconsistencies between the rigid puckered models and SANS measurements may be due to flexibility of the triskelion, since we have ruled out other possibilities such as smearing of the data from dispersion in neutron wavelengths (see Fig. 4.12) and polydispersity by DLS measurements. Analysis of SANS data with a model triskelion in which flexibility of the legs is taken into account will be discussed in the next chapter.

Secondary findings from this study involve the effect of a D_2O solvent on clathrin basket assembly in MES buffer. By DLS and a pelleting assay, it was observed that D_2O promoted basket assembly in two ways. It increased the rate of assembly at a measured pD* equivalent to the pH in an H_2O solvent (see Fig. 4.4), and increased the total amount of protein assembled after a period of 24 hours (see Table 4.2[80]).

SANS measurements of clathrin *baskets* at pH 6.0 in D_2O solvent seemed to agree well with calculations based on the D6 basket cryo-EM structure. SANS measurements of baskets formed in H_2O at the same pH appeared larger (data not shown) and did not fit the D6 basket structure. This is a subject of current study.

The fit of the rigid D6 basket model to SANS measurements of clathrin baskets in D_2O is qualitatively similar to that seen in our study of single triskelia (see Fig.

4.13). This indicates that the D6 basket structure may be a good descriptor of clathrin baskets assembled in D_2O and that clathrin baskets, too, may be flexible. This also is a subject of continued research.

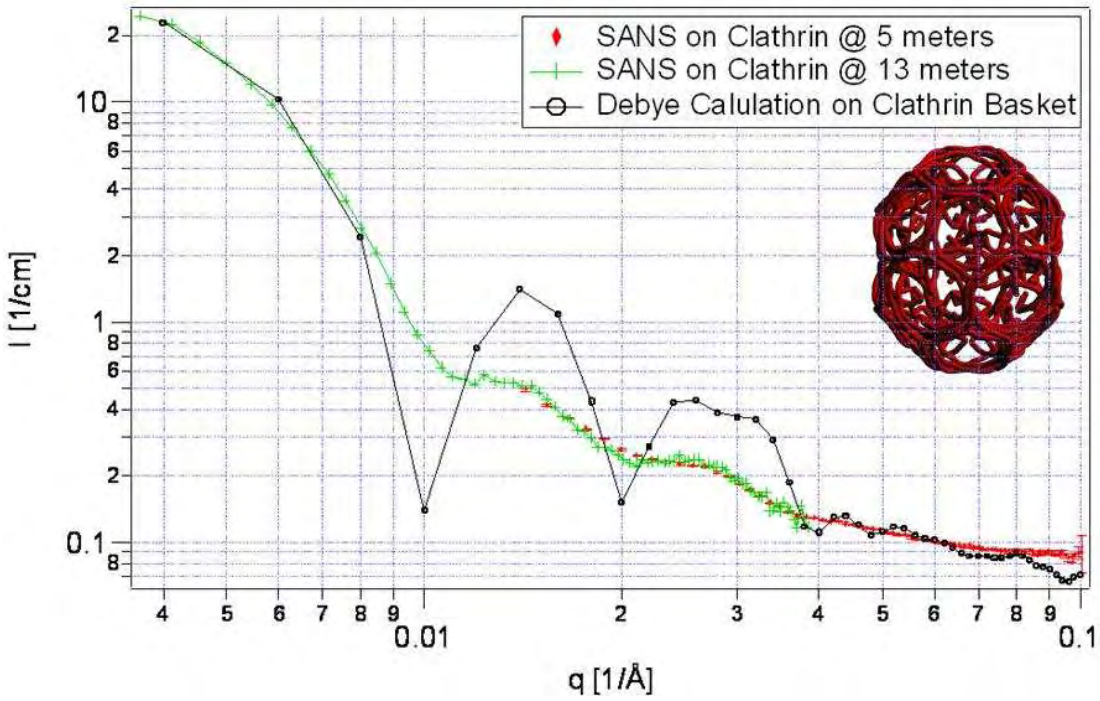

Fig. 4.13: SANS data from clathrin basket assembled in 0.1 M MES D_2O buffer at pH 6.0 and calculated scattering function based on cryo-EM model. (Shown in inset[65].) The calculated scattering function which is normalized to 1 was multiplied by a factor of $40cm^{-1}$ in order to get best agreement with SANS data. A residual incoherent background of 0.05 cm^{-1} was added to the calculated scattering function. A representation of the bead model used for the calculation is shown in the inset.

5. BROWNIAN DYNAMICS SIMULATIONS OF A BEAD-SPRING MODEL CLATHRIN TRISKELION

5.1 Introduction

In order to investigate the possible effects of thermal motions of triskelial legs on the shape of the scattering function, we constructed a discrete bead-spring model of a complete triskelion. We approximated intra-molecular forces on each subunit by linear springs. A similar bead-spring model was originally proposed by Harris and Hearst[82] to approximate semi-flexible linear polymers and later studied by Marques et al.[73]. In their model bonds between molecular subunits have identical spring constants and bond angles are preserved by springs between next-nearest neighbors having a different spring constant. Unlike the worm-like chain model of Kratky and Porod[31], which uses a fixed bond length, the Harris and Hearst model allows for bond length fluctuations determined by the strength of the springs and the temperature of the reservoir. In our model, equilibrium bond lengths are determined from a structural model and bond length fluctuations are suppressed by the use of stiff springs. This will be discussed more thoroughly in Section 5.3. The details of our model serve to preserve bond lengths (through nearest neighbor springs) and bond angles (through next-nearest neighbor springs) while allowing flexibility of the triskelion legs. Bead

size and equilibrium configurations were chosen according to the models described in Chapter 3, which were based on structural and light scattering measurements of clathrin triskelia (see Fig. 3.5). The models consistent with the measured R_H and R_g are shown in Fig. 3.10 a-d.

We simulated dynamics of the triskelion model in a thermal bath by integrating a Langevin equation[83] of the form

$$m_i \frac{d^2\vec{x}_i}{dt^2} = \vec{\mathscr{F}}_i(t) + \vec{F}_i(t), \tag{5.1}$$

where m_i is the mass of a subunit i, $\vec{x}_i$ is the subunit position, $\vec{\mathscr{F}}_i(t)$ is the total intramolecular force on a subunit i at time t and $\vec{F}_i(t)$ is the random thermal force on the subunit from the surrounding fluid.

5.2 Methods

5.2.1 Constructing a Bead-Spring Model of a Triskelion

The mass of each subunit, $m_i = \frac{4\pi}{3}\rho a_i^3$, was calculated using the density of ρ =1.38 g/mL (0.831 Da/Å^3) consistent with experimental measurements of other proteins[84]. The radius of a subunit from the model was a_i = 15 Å for legs and 25 Å for terminal domains. The mass of each subunit of the legs was approximately 12 kDa and the mass of each terminal domain was 54 kDa. This approximated the known molecular weight of a clathrin triskelion ($M_w \sim 651\ kDa$) to within 13%.

The intramolecular force on each subunit, $\vec{\mathscr{F}}_i(t)$ was approximated by springs of strength k between two neighboring subunits and springs of strength αk between two next-nearest neighboring subunits. Each bead in the leg of a triskelion had, on

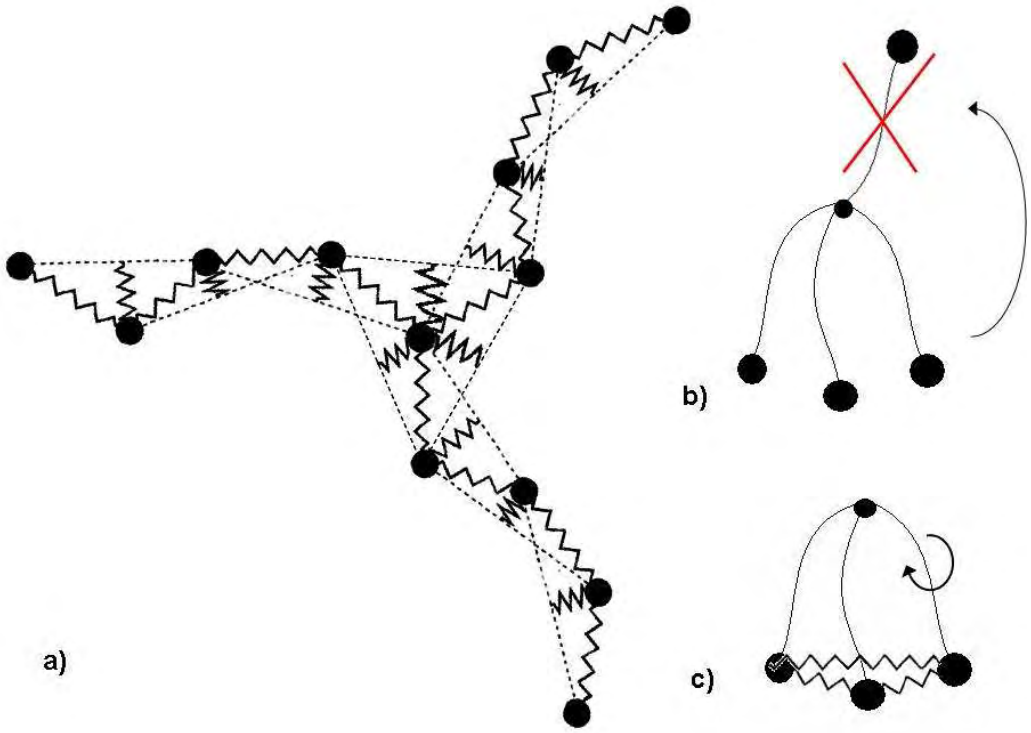

Fig. 5.1: A sketch of the bead-spring model of a clathrin triskelion used in the Brownian dynamics simulation. a) Connectivity of the vertex and legs of the triskelion model. Jagged lines represent springs between neighboring subunits and dotted lines indicate springs between non-neighboring subunits. b) Torsional motion of the legs is partially prohibited by the addition of a single spring of strength k between each termial domain. This is shown schematically in (c). c) Connections between terminal domains in the triskelion model. These springs prevent the rotation of the legs about the vertex as indicated by figure above. Some axial rotation is still unconstrained about the axis defined by the triskelion vertex and the end of the leg (shown by arrow). Figure (a) is based on a scheme introduced in Marques et al.[73]

average, four springs connected to it. Each terminal subunit, where there is only one next-nearest neighbor, had two. The vertex where there were three nearest neighbors and three next nearest neighbors was connected to other subunits by six springs. This constrained the angles between each leg at the vertex. A schematic illustration of the connections used between subunits is shown in Fig. 5.1.

In order to constrain the legs of the triskelion from twisting, each of the terminal subunits was connected by a spring of strength k (identical to the strength of springs between neighboring subinits in the legs). The legs, however, were left free to swivel about the axis defined by the vertex and terminal domain of the leg (see Fig. 5.1). This swiveling could manifest itself in a long timescale oscillation in any shape dependent quantities calculated from the simulation. This motion could be constrained by the further addition of springs in the model, but this was not done in the calculations that are shown here.

The springs constrain the motion of each subunit to fluctuate about an equilibrium structure. $\vec{\mathscr{F}}_i(t)$ is a function of the coordinates of neighboring subunits and for typical subunits can be written explicitly as

$$\begin{aligned}\vec{\mathscr{F}}_i(t) &= \alpha k(\vec{x}_{i-2}(t) - \vec{x}_i(t) - \vec{x}_{i-2}(0) + \vec{x}_i(0)) \\ &\quad + k(\vec{x}_{i-1}(t) - \vec{x}_i(t) - \vec{x}_{i-1}(0) + \vec{x}_i(0)) \\ &\quad + k(\vec{x}_{i+1}(t) - \vec{x}_i(t) - \vec{x}_{i+1}(0) + \vec{x}_i(0)) \\ &\quad + \alpha k(\vec{x}_{i+2}(t) - \vec{x}_i(t) - \vec{x}_{i+2}(0) + \vec{x}_i(0)) \qquad (5.2)\end{aligned}$$

The initial position of each subunit, $\vec{x}_i(0)$, was chosen from the models shown in Fig. 3.10. These are four rigid bead models that have the same hydrodynamic radius,

r_H, and radius of gyration, r_g, as those measured by light scattering experiments performed on clathrin triskelia in solution (see Chapter 3).

5.2.2 Estimating Model Parameters from Experimental Measurements of Clathrin Triskelia

The leg length distribution[9] and the persistence length[10] of the legs of real clathrin triskelia have been estimated from the analysis of variations in electron micrographs of clathrin triskelia flattened on a mica surface. These estimates for the elasticity of a triskelion leg will be used to choose the model parameters, k and α, of our bead-spring model of a triskelion.

The Stretching Elasticity Estimated From a Distribution of Leg Lengths

The variability in triskelion leg lengths observed in electron micrographs can give us an estimate of the longitudinal elasticity of a triskelion leg. In the study of Kocsis et al.[9], 357 randomly chosen triskelion legs were digitized and digitally straightened. The resulting data showed a Gaussian distribution of leg lengths (see Fig. 5.2). The leg lengths, L, in the micrographs analyzed varied from $350 - 620$ $\mathring{A}$ with a mean of $< L >= 516$ $\mathring{A}$ and a standard deviation of $\sigma_L = 43$ $\mathring{A}$.

In a continuous elastic medium, the force required to stretch an elastic material of cross sectional area, A, is related to the strain $\frac{\Delta L}{L}$ by the Young's modulus, E of the material[85],

$$F = -EA\frac{\Delta L}{L} \tag{5.3}$$

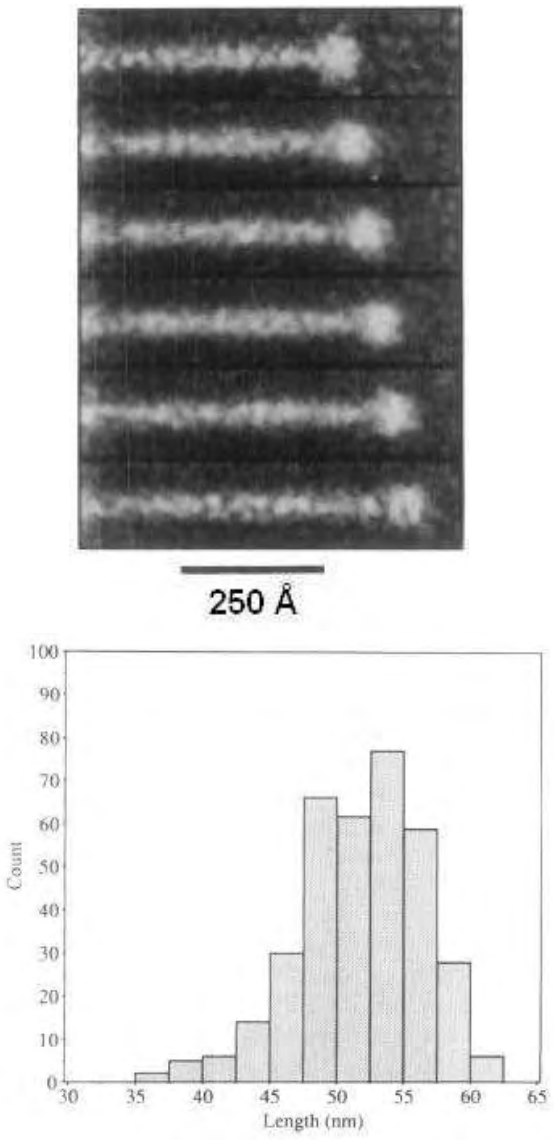

Fig. 5.2: Experimental variability in lengths of digitally straightened triskelion legs. (Top) Reconstructed images of digitally straightened triskelion legs. (Bottom) Histogram of leg lengths from 357 triskelion legs. The leg lengths in the micrographs analyzed varied from $350 - 620$ $\mathring{A}$ with a mean of $< L >= 516$ $\mathring{A}$ and a standard deviation of $\sigma_L = \sqrt{< (\Delta L)^2 >} = 43$ $\mathring{A}$. Figures 2 and 3 from Kocsis et al.[9]

and the work, W, required to extend it is given by

$$W = \int_0^{\Delta L} \frac{EA}{L} x dx = \frac{1}{2}\frac{EA}{L}(\Delta L)^2.$$

In equilibrium, the average energy stored in a fluctuating leg can be related to the temperature by the equipartion theorem:

$$< W >= \frac{1}{2}\frac{EA}{L} < (\Delta L)^2 >= \frac{1}{2}k_B T.$$

This analysis gives us an estimate for the quantity $\frac{EA}{L}$ from the leg length distribution measured by Kocsis et al.,

$$\frac{EA}{L} = \frac{k_B T}{< (\Delta L)^2 >} \sim \frac{240\ Da\mathring{A}^2/ps^2}{(43\ \mathring{A})^2} \sim 0.13\ Da/ps^2. \tag{5.4}$$

In a discrete system of masses connected by springs with equilibrium separation, b, and strength, k, the force related to an extension, δx, is given by

$$F = -k\delta x = -kb\frac{\delta x}{b}$$

. By analogy to Eq. 5.3, we can see that the quantity kb is the discrete analog to the continuous quantity EA. This allows us to estimate a value for k if the entire leg were made up of a single spring of equilibrium length $b = L$.

$$k = \frac{EA}{L} \sim 0.13\ Da/ps^2. \tag{5.5}$$

Since each leg in our model consists of 18 subunits (including the vertex and terminal domain) and since each molecular subunit is connected to two springs on average, one to its neighbor of strength, k, and one to its next nearest neighbor of strength, αk, this network of springs results in a nontrivial relationship between the

model parameters k and α and the stretching elasticity of our bead-spring model. This relationship can be investigated by the analysis of the Brownian dynamics simulation and will be discussed more in Section 5.3.

The Bending Rigidity of a Triskelion Leg Estimated from an Experimental Distribution of Leg Shapes

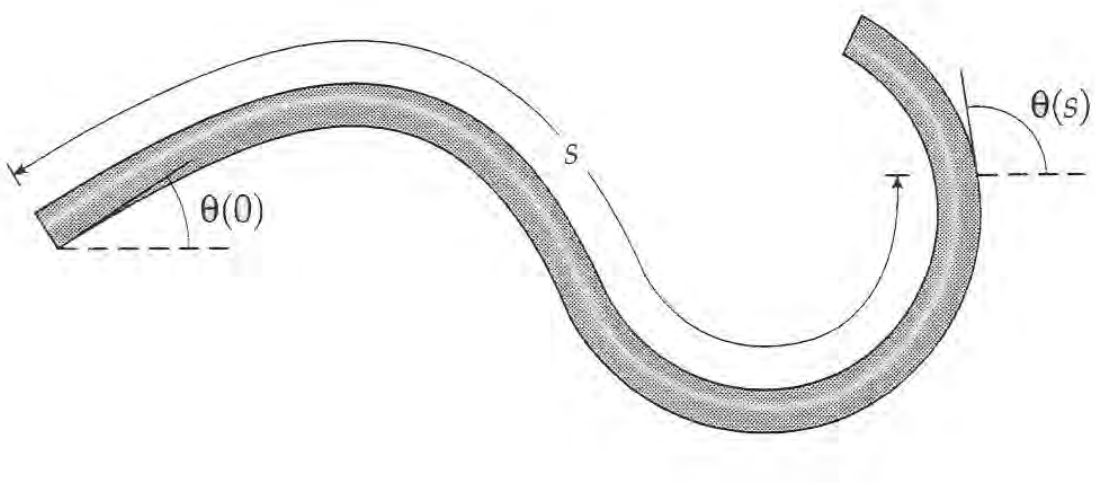

Fig. 5.3: "Diagram of an elastic filament defining the tangent vector in two dimensions. As the arc length increases, the tangent angles, $\theta(s)$ and $\theta(0)$ become uncorrelated." Figure and caption from Howard, 2001[85, Fig. 6.8].

Fluctuations in the shape of a linear filament can be characterized by a quantity called the persistence length, L_p. The persistence length is a measure of the flexibility of a filament and can be determined experimentally from the spatial correlation function of the tangent vector, $\hat{\theta}(s)$, along the filament where s is the arc-length measured along the filament($\hat{\theta}$ and s are shown schematically in Fig. 5.3). The spatial correlation function of the cosine between two tangent vectors, $\hat{\theta}(0)$ and $\hat{\theta}(s)$, at a dis-

tance s from each other along the filament decays exponentially in three dimensions according to

$$< \hat{\theta}(s) \cdot \hat{\theta}(0) >= e^{-s/L_p} \tag{5.6}$$

where $< \cdot >$ denotes a time or ensemble average. Some examples of fluctuating filaments of varying flexibilities, rigid ($L_p \gg L$), semiflexible ($L_p \sim L$), and flexible ($L_p \ll L$), are shown in Fig. 5.4.

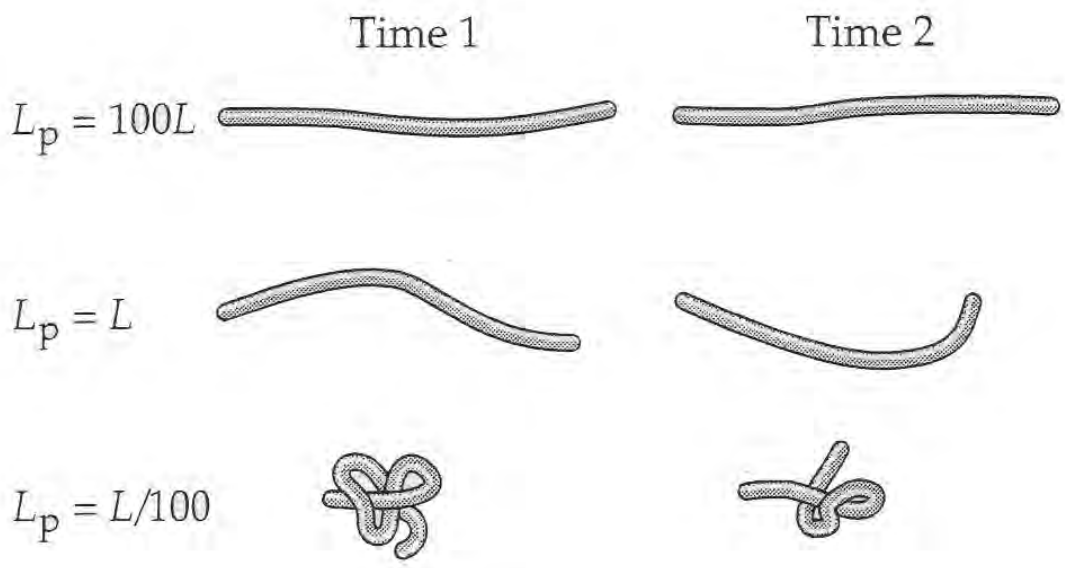

Fig. 5.4: "Two snapshots of slender rods of length L bent by thermal forces. The more flexible the rod (the smaller the persistence length), the greater the curvature." Figure and caption from Howard, 2001[85, Fig. 6.7]

In a continuous elastic medium of Young's modulus, E, and cross sectional second moment, $I = \int r^2 dA$, the persistence length of an elastic filament can be related to the flexural rigidity, EI, through the equipartition theorem. In two dimensions, the energy stored in the deformation of an elastic filament of length, L and the tangent

angle $\theta(s) = \sqrt{\frac{2}{L}} \sum_{n=0}^{\infty} a_n \cos\left(\frac{n\pi x}{L}\right)$ is given by

$$U = \frac{1}{2} EI \sum_{n=1}^{\infty} \left(\frac{n\pi}{L}\right)^2 (a_n - a_n^0)^2$$

where a_n^0 is the Fourier mode a_n in the absence of forces[86, 10, 85].

In equilibrium the average energy stored in a Fourier mode, n, is $\frac{1}{2}k_BT$:

$$< U >= \frac{1}{2} EI \left(\frac{n\pi}{L}\right)^2 \left\langle (a_n - a_n^0)^2 \right\rangle = \frac{1}{2} k_B T.$$

The flexural rigidity can be related to persistence length, L_p, of a linear filament by

$$L_p = \frac{EI}{k_B T}.$$

The shape fluctuations in electron micrographs of triskelion legs can give us an estimate of the bending rigidity of a triskelion leg. In the study of Jin and Nossal[10], 72 digitized electron micrographs of triskelion legs were analyzed to determine the bending rigidity of a triskelion leg. Some examples of these micrographs and digital triskelia are shown in Fig. 5.5.

Using this analysis, Jin and Nossal estimated the flexural rigidity of a triskelion leg to be $EI = 350\ k_BT\ \mathring{A}$ using the first 5 Fourier modes. The analysis gives triskelion legs a persistence length of $350\mathring{A}$ or $\frac{\xi}{<L>} = \frac{350\mathring{A}}{516\mathring{A}} \sim \frac{2}{3}$ where $< L >$ is the average length of a triskelial leg. The persistence length, along with the estimate of stretching flexibility, $\sqrt{(\Delta L)^2} \sim 43\mathring{A}$ or $\frac{\sqrt{(\Delta L)^2}}{<L>} \sim \frac{43\mathring{A}}{516\mathring{A}} \sim 0.08$, can be directly compared with calculations from the simulation of our bead-spring model of a triskelion to choose appropriate values of the model parameters k and α.

Similar to the stretching elasticity, our bead-spring model, has a complex relationship of the model parameters k and α to the persistence length. This dependence

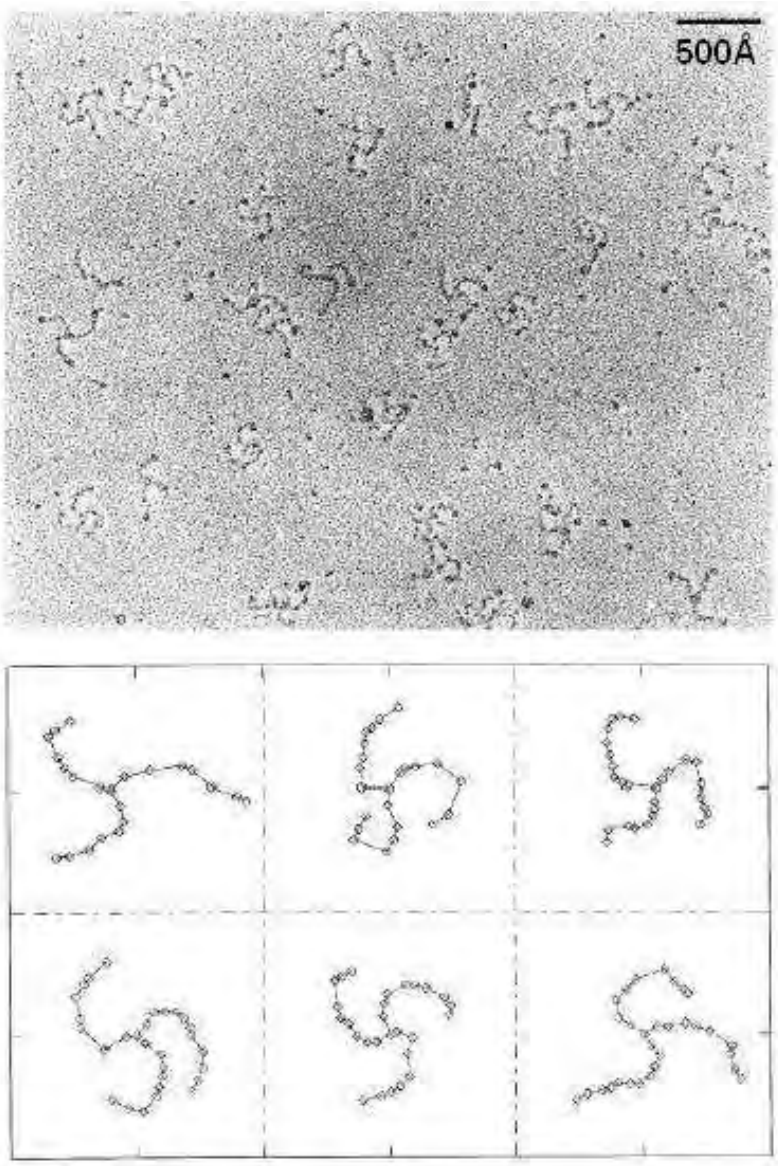

Fig. 5.5: "Experimental variability in shapes of triskelia on mica . (Top) Electron micrographs of triskelia on a mica surface. (Bottom) selected triskelion images, constructed from corresponding digitized coordinates." The persistence length was estimated by normal mode analysis to be $L_p \sim 350$ $\mathring{A}$. Figure and Caption from Jin et al.[10]

will be investigated by analysis of results from the Brownian dynamics simulation and discussed more in Section 5.3.

5.2.3 Taking into Account the Solvent

The force on an intra-molecular subunit due to the solvent, $F_i(t)$, in Eq. 5.1 may be split into a slowly varying frictional force proportional to the velocity of the subunit, $\zeta_i \frac{d\vec{x}_i}{dt}$, and a rapidly varying force, $F_i'(t)$, that averages to zero

$$m_i \frac{d^2\vec{x}_i}{dt^2} = \vec{\mathscr{F}}_i(t) - \zeta_i \frac{d\vec{x}_i}{dt} + \vec{F}_i'(t); \left\langle \vec{F}_i'(t) \right\rangle = 0, \tag{5.7}$$

where ζ_i is the friction constant of a single subunit i. The friction constant on molecular subunits was approximated by Stokes' law $\zeta_i = 6\pi\eta a_i$ where η is the viscosity of the surrounding fluid ($\eta_{Water} = 0.01 Pas = 602 \frac{Da}{\text{Å } ps}$) and a_i is the radius of the spherical subunit[87].

5.2.4 Integrating the Equations of Motion

New coordinates for each subunit of our molecule are determined by a simple Eulerian integration of Eq. 5.7 in the diffusive limit, where $\frac{\zeta}{m}\Delta t \gg 1$ and the inertial term, $m_i \frac{d^2\vec{x}_i}{dt^2}$, is small and can be ignored during integration[88, 89]. Explicitly, each subunit position was evolved according to the algorithm

$$x_{n+1} = x_n + \frac{\Delta t}{\zeta} \mathscr{F}(\{x_n\}) + X_n \tag{5.8}$$

where x_{n+1} is the new subunit position, x_n is the old subunit position, $\mathscr{F}(\{x_n\})$ is the total intramolecular force on a subunit at time step n, Δt is the time step and X_n is a random displacement simulating thermal collisions with solvent molecules.

5.2.5 Approximating the Temperature

For the random variable X_n in Eq. 5.8, a uniform random distribution, $\rho(X_n)$, was used[88, 90]. The variance in X_n determines the temperature of the solvent in the simulation. Explicitly, X_n is chosen from a uniform distrubution between $\pm\sqrt{6D_i\Delta t}$, where $D_i = \frac{k_B T}{\zeta_i}$ is the diffusion coefficient of the subunit i in the absence of other subunits. This gives the second moment of X_n a value of $2D_i\Delta t$, shown below.

$$< X_n^2 >= \frac{\int_{-a}^{a} \rho(X_n) X_n^2 dX_n}{\int_{-a}^{a} \rho(X_n) dX_n} = \frac{1}{3}a^2 = 2\frac{k_B T}{\zeta_i}\Delta t = 2D_i\Delta t$$

where,

$$\begin{aligned} \rho(X_n) &= 1,\ |X_n| < a \\ &= 0,\ o\mathit{therwise}. \end{aligned}$$

We ignored hydrodynamic interactions between subunits in order to facilitate the simulation. Incorporating hydrodynamic interactions between molecular subunits in the dynamical simulation would have the affect of reducing the friction coefficient on each subunit by hydrodynamic shielding from other parts of the molecule. Therefore any diffusion coefficient calculated from the mean squared displacement of the molecule in the simulation would be too large. Structural dependent quantities calculated from the simulation might also be affected.

We are interested in fluctuations in the shape of the triskelion molecule. Since we are underestimating the frictional coefficients of the subunits, we should consider that the conformational fluctuations calculated from these simulations may be overly dampened. Since the simulations are performed in the Brownian dynamics limit, the friction coefficient scales out of the equation and may only effect the time scales of conformational fluctuations and not the overall magnitude of fluctuation. The most significant affect of hydrodynamic interactions may be the coupling and consequent correlation of the subunit positions that occur because of this term. The effect of hydrodynamic interactions on Brownian dynamics and other stochastic simulations is discussed thoroughly by McCammon et al.[91] and Pastor et al.[89, 87].

Choosing a Timestep (Estimating Numerical Error)

The numerical integration of Eq. 5.7 has an error associated with it that is dependent upon the time step Δt used in Eq. 5.8. We can evaluate this error for the simple case of a particle in a harmonic potential of strength k and a thermal reservoir at temperature T[87]. The mean squared displacement $< x^2 >$ is given by $\frac{k_B T}{k}$. Squaring Eq. 5.8 and taking the time average, denoted by $\langle \cdot \rangle$ we get

$$\begin{aligned} < x_{n+1}^2 > &= \left\langle \left(x_n - \frac{k\Delta t}{\zeta} x_n + X_n \right)^2 \right\rangle \\ &= \left(1 - \frac{k\Delta t}{\zeta} \right)^2 \left\langle x_n^2 \right\rangle + \left(1 - \frac{k\Delta t}{\zeta} \right) \langle x_n X_n \rangle + \left\langle X_n^2 \right\rangle . \end{aligned}$$

We now assume that $< x_{n+1}^2 >=< x_n^2 >$ and $\langle x_n X_n \rangle = 0$,

$$\begin{aligned} 0 &= -\frac{2k\Delta t}{\zeta} \left(1 - \frac{k\Delta t}{2\zeta} \right) \left\langle x_n^2 \right\rangle + \frac{2k_B T}{\zeta} \Delta t, \\ \left\langle x_n^2 \right\rangle &= \frac{k_B T}{k} \frac{1}{1-\epsilon}; \; \epsilon = \frac{k\Delta t}{2\zeta}. \end{aligned} \tag{5.9}$$

The time step, Δt, was chosen such that the numerical error, ϵ, in $\langle x_n^2 \rangle$ of the strongest spring used in the simulation, k, was 5%, or explicitly that $\Delta t = \frac{1}{10}\frac{\zeta}{k}$.

Checking the Bead-Spring Model

In order to be certain that the forces were implemented correctly, the intramolecular force, $\mathscr{F}(\{x_n\})$ was checked for energy conservation using the Verlet algorithm. This is a well known molecular dynamics algorithm used to integrate the dynamics of Hamiltonian systems[92, 87]. The Verlet algorithm has no dissipative or friction term and is not connected to a thermal bath. Explicitly, the new positions determined by the Verlet algorithm are calculated from previous ones according to

$$x_{n+1} = 2x_n - x_{n-1} + \frac{\Delta t}{m}\mathscr{F}(\{x_n\})$$

Subunits were initialized with a random velocity chosen from a uniform distribution between $\pm\sqrt{\frac{3k_BT}{m_i}}$, our approximation of a Boltzman distribution. The energy as a function of time is shown is Fig. 5.6. While the total energy fluctuates about an average value, determined by initial velocities of each subunit, the average total energy remains constant over both short (ns) and long timescales (μs).

The Computation of Time Averages from Dynamic Simulations

In order to determine how long to run the simulation for computation of a meaningful time average of dynamic quantities in the simulation, the timescales related to internal modes of oscillation of the bead-spring model can be estimated and any time average should be computed using a simulation that is run for intervals that are longer than this timescale.

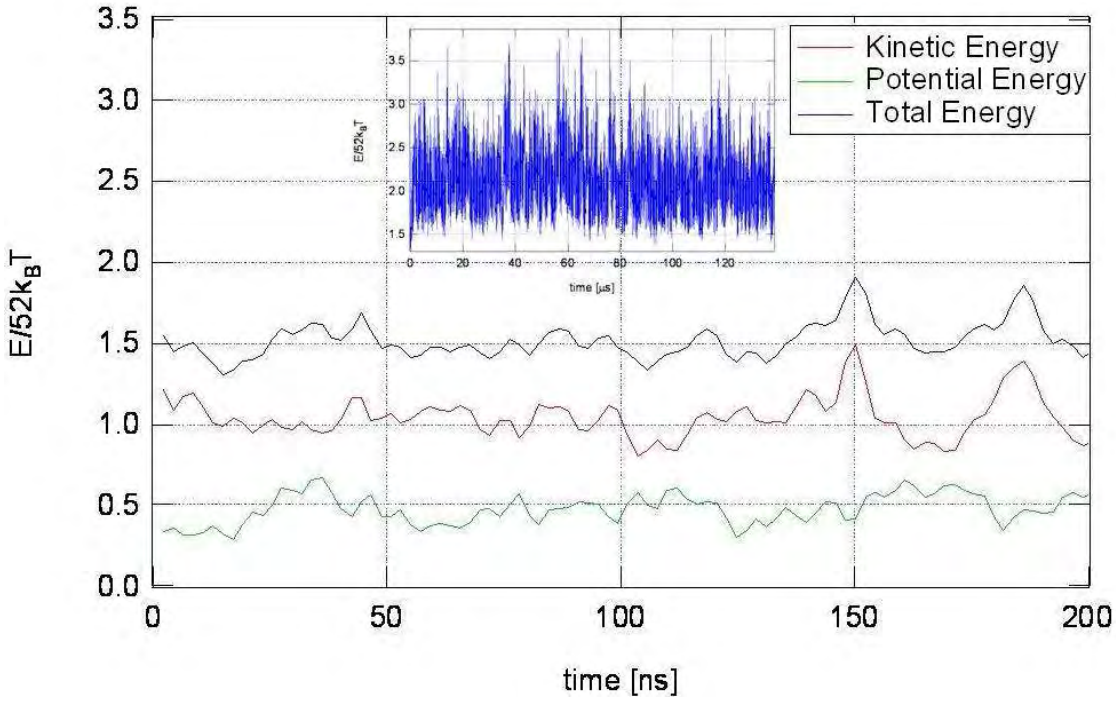

Fig. 5.6: Energy conservation of a molecular dynamics simulation of a triskelion. This figure shows average total energy, kinetic energy and potential energy per subunit per k_BT vs. simulation time for a triskelion model using the molecular dynamics Verlet algorithm. The initial velocities of subunits were chosen randomly from the interval $\{-\sqrt{\frac{3k_BT}{m_i}}, +\sqrt{\frac{3k_BT}{m_i}}\}$. Subunits were started in thier equilibrium positions, so the total initial potential energy is zero. In this example the model parameters were $k = 100\ Da/ps^2$ and $\alpha = 10$. The timestep was 0.705 ps. Snapshots were taken at 2.1 ns intervals.

Since we are interested in computing a time averaged scattering function, we will concentrate on fluctuations in the radius of gyration which is loosely related to the width of the central peak in the scattering function. The radius of gyration can be defined as the second moment of the mass distribution in a molecule. If $\rho(\vec{r})$ is the mass (or scattering) distribution of a molecule, r_g can be calculated for a bead-spring model according to the equation

$$r_g = \int_V d\vec{r}|\vec{r}|^2\rho(\vec{r})$$

If the molecule is modeled as a discrete set of spherical subunits where $\vec{r}_i$ is a subunit position with respect to the center of mass, a_i is the subunit radius, and $v_i = \frac{4\pi}{3}a_i^3$ is a subunit volume, r_g can be calculated according to the simple sum

$$r_g = \frac{1}{N}\frac{\sum_{i=1}^{N} v_i|\vec{x}_i|^2}{\sum_{i=1}^{N} v_i}$$

The time dependent radius of gyration, $r_g(t) = r_g(\{\vec{x}_i(t), a_i\})$ of the bead-spring triskelion model can be considered a dynamic variable and is a function of the bead coordinates and radii. The temporal autocorrelation function of the instantaneous radius of gyration can tell us something about the timescales with which the model conformation, and therefore the scattering function, is fluctuating.

In the Brownian dynamics simulation, r_g if found to fluctuate about an average value $\frac{(<r_g>-r_{g0})}{r_{g0}} = 5\% - 18\%$ below the value of the r_g for the rigid model, r_{g0} and approaches the value, r_{g0} as the spring strength is increased (in the rigid model limit). The RMS fluctuation in the r_g is on the order of 10 Å and is a weak function of spring strengths, k, and αk. The time averaged quantities $< r_g >$ and σ_{r_g} for a few different spring strengths are shown in Fig. 5.8.

The temporal autocorrelation function of r_g, denoted $\delta r_g = r_g(t) - < r_g >$ was calculated by the following equation,

$$\frac{< \delta r_g(\tau)\delta r_g(0) >}{< \delta r_g^2 >} = \frac{\frac{1}{T}\int_0^T \delta r_g(t+\tau)\delta r_g(t)dt}{\frac{1}{T}\int_0^T \delta r_g(t)^2 dt} \sim e^{-\tau/\tau_c}$$

and had an initial, rapid exponential decay in delay time, τ_c, between 5 and 13μs, for different spring strengths, k. A few correlation functions are shown in Fig. 5.7 along with their exponential fits.

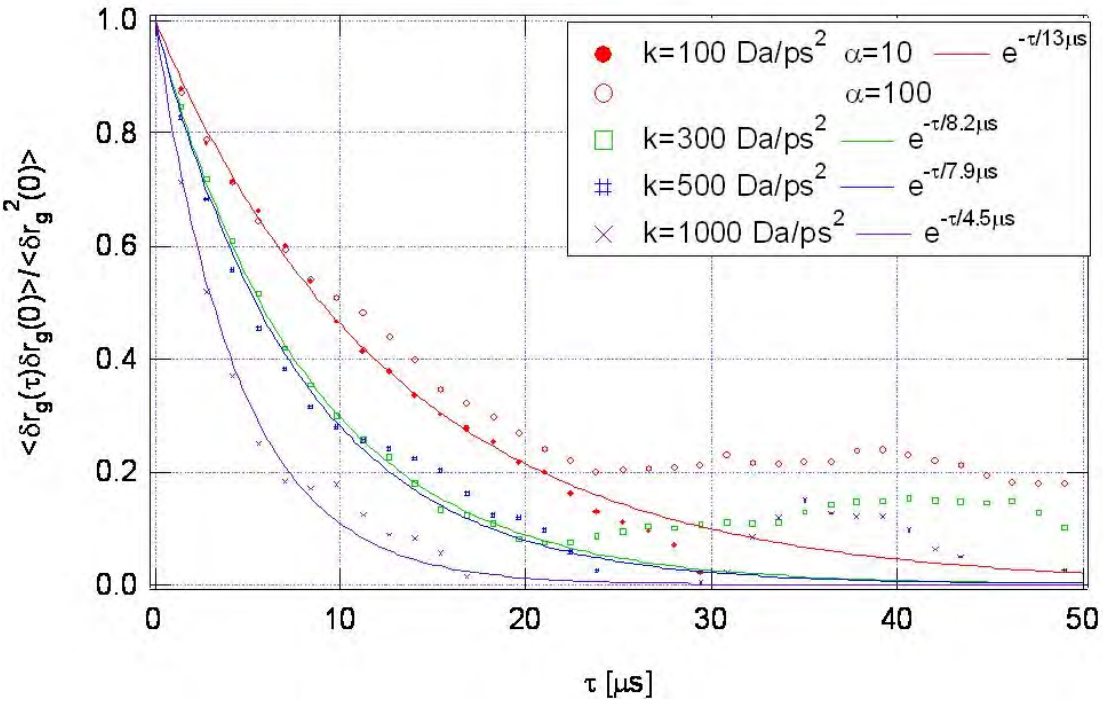

Fig. 5.7: The temporal autocorrelation function of r_g decays with a timescale on the order of 10 μs. The correlation time showed some sensitivity to the spring strength, k, of the triskelion model but did not seem to be sensitively dependent on the model parameter α. Brownian Dynamics Simulations were run for a duration of $\sim$ 700 μs. 500 snapshots were taken at 1.4 μs intervals.

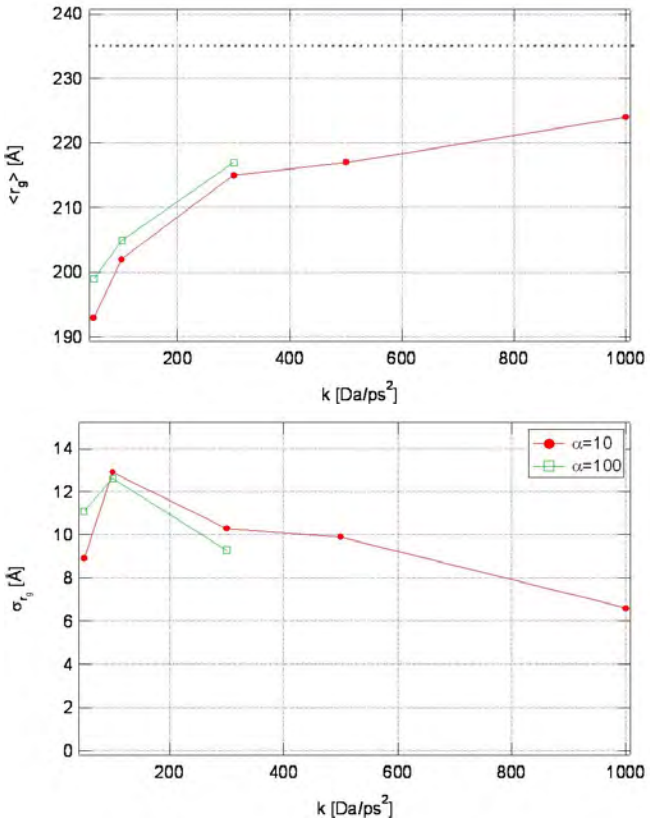

Fig. 5.8: Variation of average radius of gyration, $< r_g >$ and standard deviation $\sigma_{r_g} = \sqrt{< (\delta r_g)^2 >}$ with spring strength, k. (Top) variation in the time average of the r_g with spring constant, k. The symbols show time averages from simulation for two different values of bond angle spring strength parameter, α. The dotted line shows the r_g of the rigid model, $r_{g0} = 235\mathring{A}$, for the model calculated. (Rigid model is shown in Fig. 3.10a.) The time average, $< r_g >$, approaches this value as the spring strength, k is increased. (Bottom) the standard deviation of r_g, denoted σ_{r_g} vs. the spring strength, k. The standard deviation, σ_{r_g}, should approach 0 in the rigid model limit, but may have a fixed value due to swivel motion of the legs as previously discussed in Section 5.2 (shown schematically by the arrow in Fig. 5.1 c).

5.3 Results

The Brownian dynamics simulation was run for several different values of k and α at a temperature of 20 $^{\circ}C$ for a total duration of 700 μs. Computation time varied with time step which in turn was dependent upon the strongest spring in the simulation (αk) and the desired degree of accuracy ($\epsilon = \frac{\alpha k \Delta t}{2\zeta} = 5\%$). 500 snapshots of the configuration of each molecule were recorded durring the simulation.

For a given set of spring strengths, k and αk, the leg extension flexibility of the bead-spring model was determined by the explicit calculation of temporal fluctuations in leg lengths during the simulation. Triskelial leg length fluctuations were dependent on the spring constants used, but were less than 1% of the total leg length for the parameter range that was simulated.

The persistence length of the triskelion legs, L_p, was also found to be dependent on the spring constants used in the simulation. This parameter was determined by calculation of the spatial correlation function of the directional tangent vector cosine according to Eq. 5.6. The tangent vector cosine was calculated, as a function of arc-length s, with respect to a tangent vector defined between the vertex and the first subunit of each individual triskelion leg. Examples of the calculated spatial correlation functions for four different sets of spring constants are shown in Fig. 5.9 on a semi-logarithmic plot. One can clearly see the straight line in these semi-logarithmic plots indicating exponential decay in the correlation function. By the difference in the slopes of the data one can also see the dependence of the persistence length on the spring constant, k, and its independence of the parameter α for the two values

studied. Once calculated, the correlation functions were fit to an exponential curve with the decay length (the inverse slope of the straight line) giving a measure of the persistence length, L_p. The persistence length varied between 200 and 1000 Å for the range of model parameters studied. This can be compared with the experimentally estimated value of $L_p = 350$ Åwhich was discussed in Section 5.2.2 and determined by the analysis if electron micrographs (see Fig. 5.5).

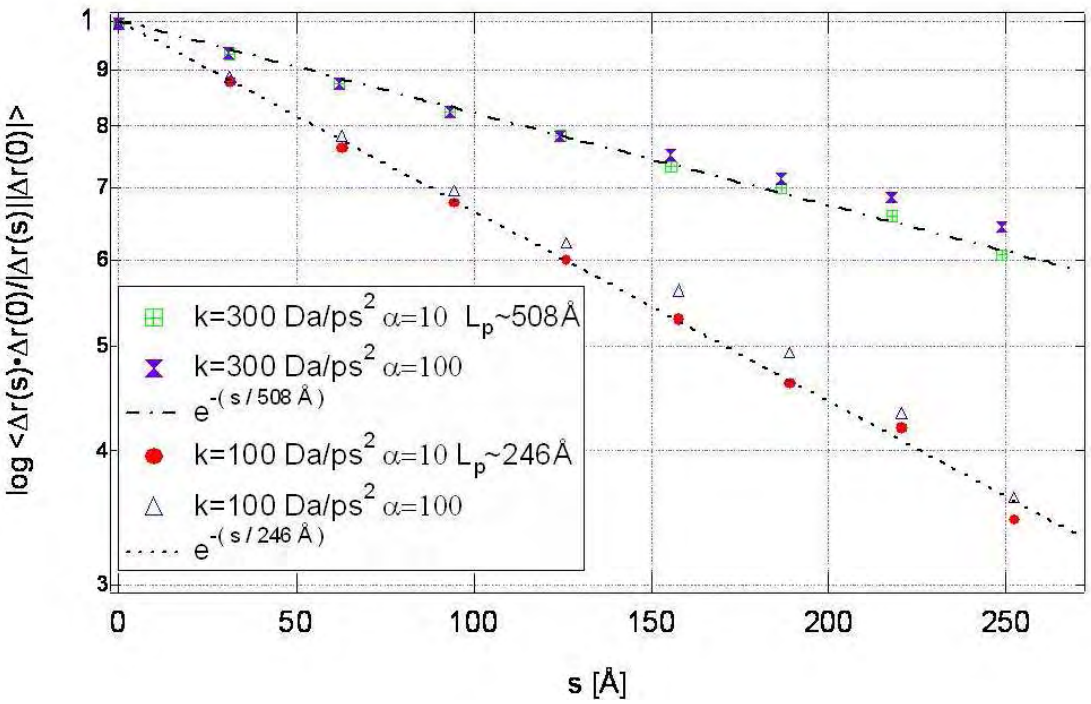

Fig. 5.9: The time average of directional correlation of tangent vectors, $< cos(\Delta\theta) >=< \frac{\Delta\vec{x}(0)\cdot\Delta\vec{x}(s)}{|\Delta\vec{x}(0)||\Delta\vec{x}(s)|} >$, between subunits in a Brownian dynamics simulation of a bead-spring model of a clathrin triskelion shows decay with position, s. This decay defines a persistence length, L_p, of the simulated triskelion leg. Symbols show time averages from the simulations. The dotted lines show exponential fits to simulations for $\alpha = 10$. $\Delta\theta$ is the angle between tangent vectors, $\Delta\vec{x}$, at position at leg position the vertex, 0, and at position, s.

In Fig. 5.10 the calculated persistence length of the simulation is plotted as a function of spring strength, k, for two different values of the model parameter α. The persistence length has an inverse square root dependence on the spring strength, k, as shown by the dotted line fit. The experimentally estimated value of the persistence length is shown by the dark black line drawn at a value of 350 Å. This corresponds to a Brownian dynamics simulation with a spring constant of $k \approx 200$ Da ps^{-2}, which has not been performed to date. The persistence length seems to be independent of α for the two values simulated.

Snapshots from the simulation were used to calculate a series of instantaneous scattering functions using Eq. 2.20. For each simulation, these functions were calculated and then averaged over time to produce an average scattering function which can then be compared with experimental measurements. Four average calculated scattering functions with different triskelion flexibilities can be seen in Fig. 5.11 along with SLS and SANS data taken on clathrin triskelia in 0.5M Tris at pH 7.0 in H_2O.

The time averaged scattering function agrees with experimental SANS data from isolated clathrin triskelia in solution for high, low and intermediate q-values where the rigid model was insufficient (see Fig. 4.6 and 4.7). This shows that, while rigid modeling was sufficient to explain average properties such as translational diffusion coefficients and the radius of gyration (see Chapter 3), more information is included in the complete scattering function. Extra information is contained in the SANS data, and the measurements allow us to say something about the molecular flexibility and dynamics of isolated triskelia.

As the rigidity of the bead-spring model is decreased, the dip and subsequent

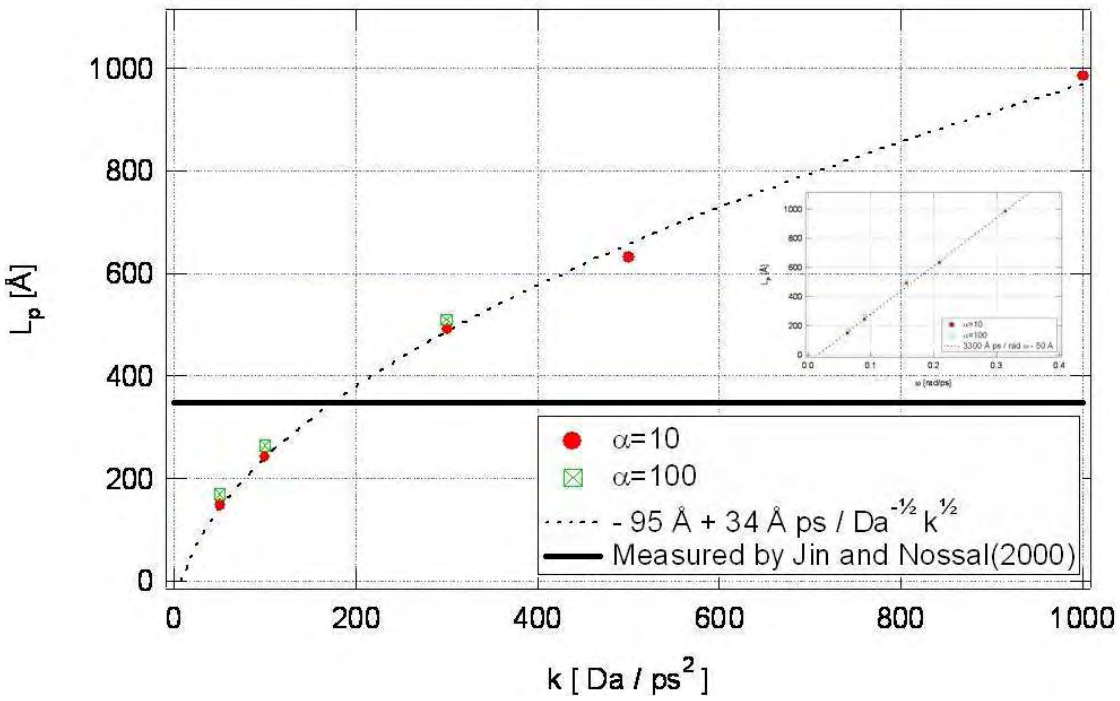

Fig. 5.10: The persistence length, L_p, depends upon the spring strength, k in a Brownian dynamics simulation of a bead-spring model performed at room temperature. Symbols show persistence lengths derived from the simulations. The dotted line shows a fit to the symbols for $\alpha = 10$. The inset shows L_p vs. frequency of oscillation of a spring for a typical leg subunit of mass, 12 kDa. L_p is a linear function of the frequency of oscillation, ω.

peak in the scattering function at intermediate q-values is smoothed out. The dip and peak in the time averaged scattering function becomes lower in amplitude eventually becoming only a shoulder-like feature. It should be noted that this shoulder occurs at the same value of q in which a shoulder is seen in the experimental data. If the flexibility of the model is increased even more, then the feature at intermediate values of q disappears. This very flexible model does not qualitatively agree with either the experimental SLS and SANS data. In Fig. 5.12, the same four time-averaged scattering functions with different model flexibilities are shown along with SLS and SANS data taken on clathrin triskelia in 0.5M Tris at pH 7.0 in D_2O. (A different experiment than shown in Fig. 5.11) This experiment gives a higher contrast between solute and solvent, allowing measurement of the scattering function to higher q-values at a detector distance of 5 meters as well as 13 meters. In this figure agreement between SANS data and the flexible model can be seen over the entire q-range measured which extends over two and a half orders of magnitude.

5.4 Summary and Discussion

By SANS, we have confirmed previous estimates of triskelion structure in solution as well as estimates of triskelion leg flexibility. We have shown that isolated clathrin triskelia have a puckered structure on average but undergo large fluctuations about this average puckered shape. The measured SANS scattering function suggests that clathrin triskelia have a considerable amount of flexibility when free in solution.

The flexibility of triskelion legs may be important in describing clathrin assem-

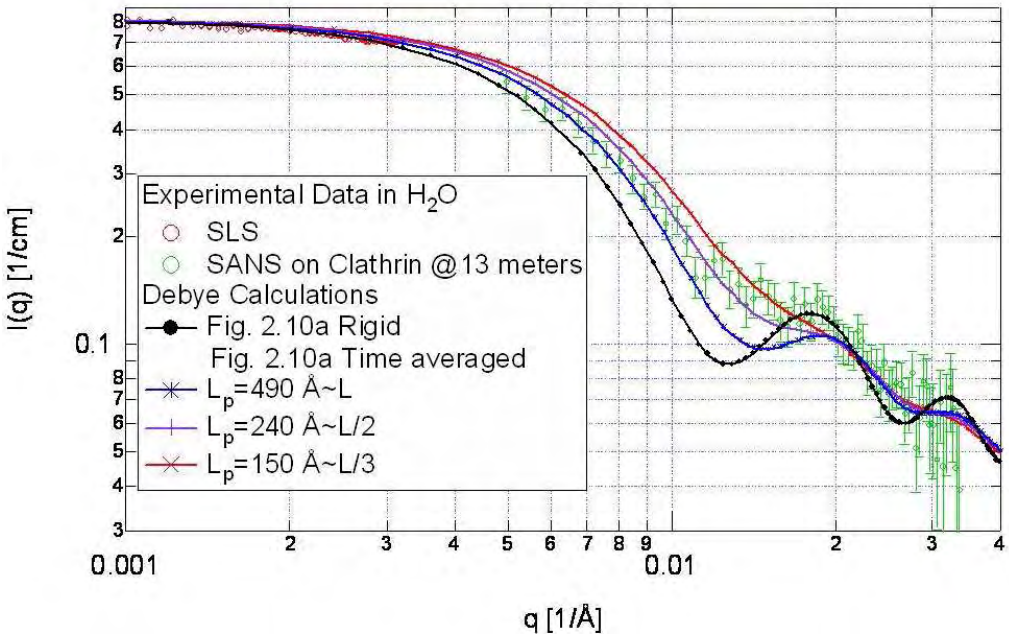

Fig. 5.11: SLS and SANS data and calculated scattering function of flexible bead-spring models of a clathrin triskelion. SLS and SANS data of clathrin in 0.5 M Tris in H_2O. A constant background of H_2O buffer was subtracted of ~ 1 cm^{-1}. The SLS data (in counts per second) were scaled to the SANS data. The line show time-averaged scattering functions calculated using Eq. 2.19, for four models of a clathrin triskelion made up of 52 beads. These are the models that fit the DLS and SLS measurements from Chapter 3. The time averages were calculated based on the Brownian dynamics simulation described in Section 5. The simulations were performed in water$\left(\eta = 602\frac{Da}{\text{Å}ps}\right)$ at a temperature of 20 $^\circ C$. The mass density used was 0.831 $\frac{Da}{\text{Å}^3}$. The spring constants used were $k = 50, 100$ and 200 Da/ps^2. The factor α was 10. The time averaged scattering function which is normalized to 1 was multiplied by a factor of 0.8 cm^{-1} in order to get best agreement with both SLS and SANS data. A residual background of 0.015 cm^{-1} was also added to the calculated scattering functions.

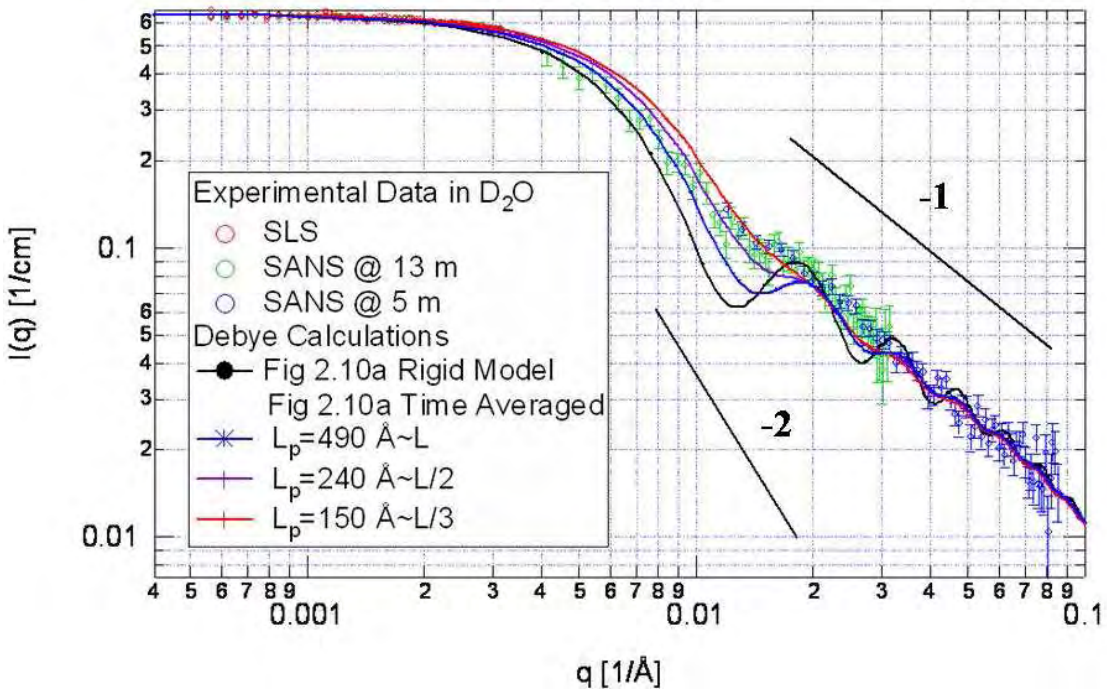

Fig. 5.12: SLS and SANS data of clathrin in D_2O and calculated scattering function of flexible bead models of clathrin. SLS and SANS data of clathrin in 0.5 M Tris in D_2O. A constant background of ~ 0.1 cm^{-1}, ascribed to the D_2O buffer, was subtracted. The SLS data (in counts per second) was scaled to the SANS data. Similar to Fig. 5.11, the lines show time averaged scattering functions calculated using Eq. 2.19 for four models of a clathrin triskelion made up of 52 beads. The data shown were taken as described in Chapter 4.

bly in many respects. It could be that triskelial flexibility is stabilized once positioned in a clathrin lattice. This would explain why structures can be determined by cryo-EM tomography and single particle analysis when these techniques are used to study a complete clathrin lattice but are not able to be rendered for isolated triskelia. If this were the case it would imply an entropic loss during clathrin lattice assembly, which may be important in basket size determination and basket stability.

6. SUMMARY, DISCUSSION, AND FUTURE WORK

6.1 Summary

Early studies by Kirchhausen et al. involved the analysis of TEM micrographs of clathrin triskelia deposited on mica surfaces. The triskelia tended to show a definite clockwise or counterclockwise bend in the three legs that depended on the buffer that was used for deposition[17]. It was argued that the different buffers modified the interactions between triskelion molecules an the mica surface, thus differentially depositing the triskelia "right side up" or "upside down." This was also confirmed by the measurement of average distances between terminal domains in the micrographs. The distance between terminal domains was larger for the "upside down" triskelia than for "right side up" triskelia. This was the first evidence that clathrin triskelia had an intrinsic, three dimensional pucker in solution. A schematic illustration from this study is shown in Figure 6.1.

Later TEM studies by Kocsis et al.[9] used digitization and particle averaging of triskelia deposited on mica to determine a leg length distribution as well as the average dimensions of different leg segments and an average two dimensional projection of the shape of single triskelia. Later analysis of these micrographs by Jin and Nossal[10] indicated that triskelion legs were flexible with a persistence length of 350

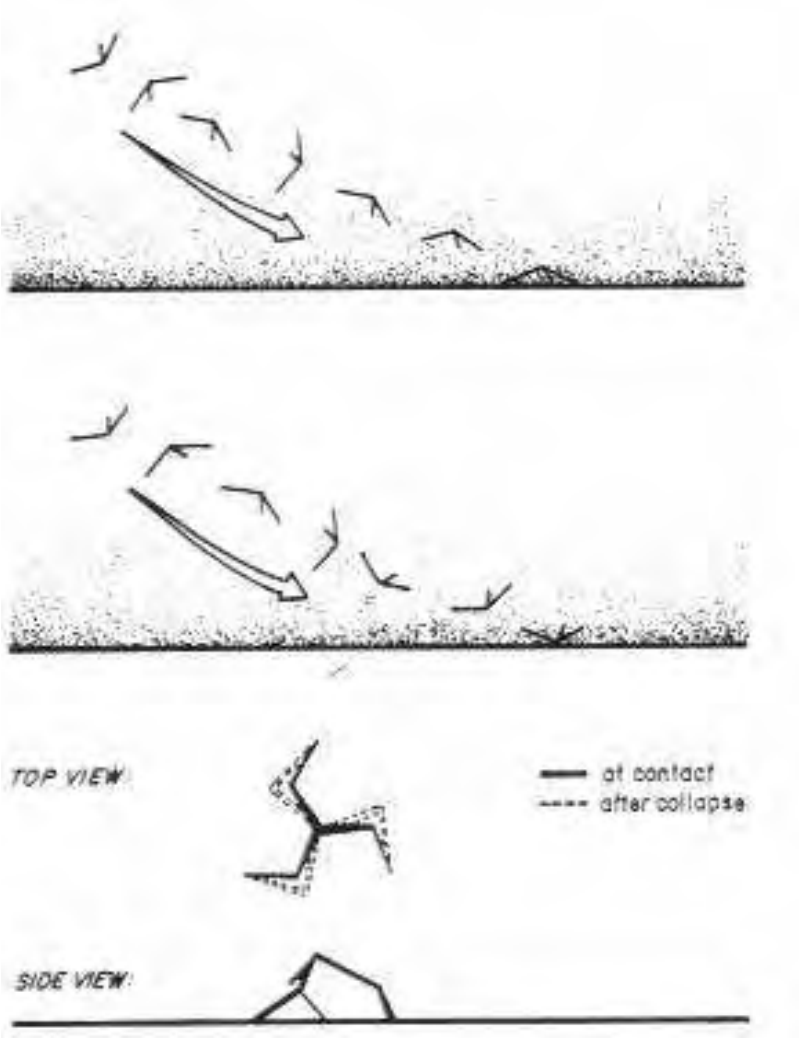

Fig. 6.1: "Schematic representation of the adsorption of clathrin to mica. Puckered clathrin triskelia are shown approaching a mica surface from their concave(top) or convex sides(middle). The bottom diagram provides top and side views of a triskelion, showing the consequence of adsorption to mica. The dotted line illustrates the outcome of the internal collapse that would result from a three-point landing. This collapse would probably include a "screwing" motion that would increase the observed bends of the proximal/distal hinges." (Figure and Caption adapted from Reference [17])

Å, approximately $\frac{2}{3}$ the length of the leg. This was done by fitting normal mode amplitudes (or spatial Fourier coefficients) to the equipartition theorem as described in Section 5.2.2

The combination of light scattering and SANS experiments described in this dissertation is consistent with previous work. They confirm both the interpretations and the conclusions of previous measurements. In the analysis of the shapes of triskelia adsorbed to carbon films or mica surfaces, details about triskelial structure and flexibility when they are in solution are inferred. However, since light scattering and neutron scattering measurements are performed while the triskelia are free in solution, they are direct, noninvasive, and nonperturbative.

The earliest analytical models of macromolecules required either complete flexibility, in the case of Gaussian chain polymeric models, or complete rigidity, in the case of the spherical, rod like, or ellipsoidal models. Models for semiflexible chains have been investigated such as the Kratky and Porod chain[31] and the early Harris and Hearst models[82]. These models have been compared with polymers and other linear molecules using light scattering, neutron scattering, and other experimental techniques. More recently, the dynamical properties of molecules of more complex shape than a linear chain have been investigated combining structural information and bead-spring models by normal mode analysis and molecular dynamics simulations. In this study, we have examined a unique macromolecule that simultaneously exhibits a complicated average structure but has flexibility about that structure. Clathrin triskelia have not one single conformation but have many different conformations in solution. Light scattering measurements described in this dissertation show that the

clathrin triskelion has a puckered structure. SANS measurements demonstrate that dynamics are present in solution. This may indicate that both play a role in the assembly and disassembly of the clathrin coat during function in eukaryotic cells.

6.2 Discussion

There is a growing understanding of proteins that do not have a single conformation but can (must) assume many different shapes. It is clear that, in order to understand the relationship between a protein structure and its function, as we are trying to do for the clathrin triskelion, the unconstrained degrees of freedom (flexibility and hinges, etc.) must be determined as well. These may be even more important than the constrained degrees of freedom that relate to a low energy static structure consistent with the folded ground state of a known polypeptide sequence (the protein folding problem). Some examples may be seen in recent work describing the mechanics of the ribosome[93], the molecular motor kinesin[94], and ATP synthase[95].

Analysis techniques, such as the one used here, that can extract information about protein conformational dynamics from established measurement techniques such as small angle scattering (light, neutron or X-ray), TEM particle averaging, and cryo-EM single particle analysis, may complement newer, single molecule techniques used in AFM pulling experiments, optical trapping experiments, and forrester resonance energy tranfer(FRET). In some cases macroscopic scattering techniques may be easier to apply to the system under consideration. They may also be be less

invasive than AFM or optical tweezers, and do not require fluorescent labeling as in the case of FRET.

6.3 Future Work

We have constructed a dynamic structural model for a single clathrin triskelion to explain experimental SANS data taken from single triskelia. (See Figure 5.11) This was done because the rigid models did not fit our experimental data at intermediate values of q. Similar techniques could be used to fit SANS data from clathrin baskets assembled at low pH. Both the data and model are shown in Figure 4.13. Utilization of a dynamic model could give us an estimate for the flexibility of the clathrin cage, which has been estimated but not measured[10]. The stiffness of a clathrin coated vesicle (different from the bare, reconstituted cage) has been recently measured by AFM to be 20 times stiffer than the estimate for a basket[96]. It would be interesting to directly determine the flexibility of the clathrin basket for comparison with this measurement.

Considerable progress has been made in the last decade in understanding the mechanisms of assembly/disassembly of the clathrin lattice. Cryo-EM tomography has played a central role in this research. It has been used to visualize position and arrangement of subunits, adaptors, and chaperone proteins within the lattice (see Figure 6.2). In the presence of adaptor protein AP2, clathrin triskelia will assemble into cages of 6-2-2 symmetry. In 1999, this structure was resolved by Musacchio et al.[65] using cryo-EM tomography at a high enough resolution to distinguish in-

dividual heavy chains within the clathrin lattice. This structure, when combined with X-ray structures of crystallized clathrin fragments, produced a high resolution model of the triskelion lattice. As resolution has improved, cryo-EM tomography has confirmed the positions of the light chains along the heavy chain (Figure 6.2a) and determined the effect of the chaperone Hsp70 and cochaperone auxilin on cage structure and regions of localization within the coat (shown in Figure 6.2 c,d[16, 97, 98]). These structures illustrate the role of each subunit in the clathrin machinery, shedding light on the mechanisms of clathrin recruitment and disassembly during endocytosis and answering questions that cannot be addressed easily by the use of X-ray diffraction or other techniques.

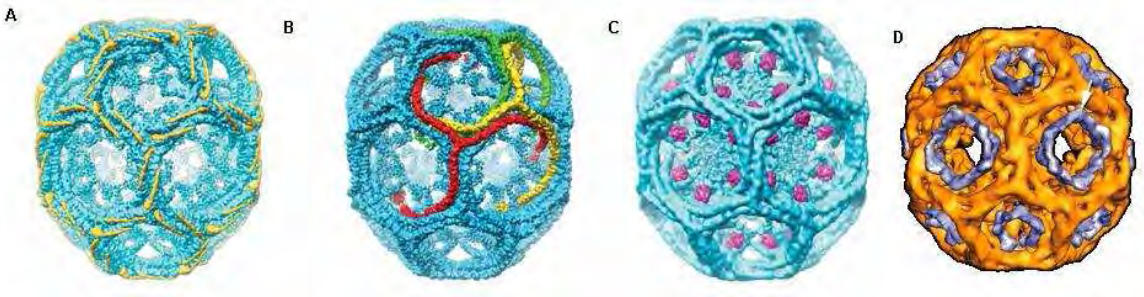

Fig. 6.2: Positions of protein components in the clathrin lattice. A) "An 8-Å reconstruction of a clathrin basket with the light chains highlighted in yellow[16]. B) An 8-Å reconstruction of a clathrin basket with individual clathrin heavy chains highlighted in red, green, and yellow." (A and B Figure and Caption from Reference[16]) C) "A 12-Å reconstruction of a clathrin basket formed in the presence of a 39-kDa fragment of auxilin (purple)." (Figure and Caption from Reference[97]) D) "A 23-Åreconstruction of a clathrin basket prepared in the presence of Hsc70 (blue)." (Figure and Caption from Reference[98])

Recently it was shown by Heymann et al.[98] that an uncoating protein, Hsp70, localizes to the polyhedral surface of an assembled clathrin basket. This protein is colored in purple in the electron density map from cryo-EM tomography shown in Figure 6.2 d. When triskelial dynamics and thermal fluctuations, which we have shown to be considerable, are taken into account, one can see that a slight modification of the equilibrium shape of a triskelion within a cage assembly (Figure 6.2 b) due to the presence of Hsp70 (or an auxilin-Hsp70 complex) might be sufficient to thermodynamically destabilize the basket, as is seen in vitro[15]. In a simulation, this mechanism for cage destabilization could be tested by shortening the equilibrium distances of springs between terminal domains of triskelia, thus altering the natural pucker of each triskelion within the lattice.

It has been noted that in the presence of the Hsp70 chaperone makes the 6-2-2 basket more rounded in shape, while the presence of the auxilin co-chaperone causes the structure to have a more barrel-like shape in 3D cryo-EM reconstructions. These structural changes could easily be confirmed by further SANS experiments.

BIBLIOGRAPHY

[1] B. Alberts, D. Bray, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter. *Essential Cell Biology.* Garland Publishing, Inc., New York, 1998.

[2] J. Fiaux, E. B. Bertelsen, A. L. Horwich, and K. Wuthrich. Nmr analysis of a 900k groel-groes complex. *Nature*, 418(6894):207–211, 2002.

[3] R. Riek, J. Fiaux, E. B. Bertelsen, A. L. Horwich, and K. Wuthrich. Solution nmr techniques for large molecular and supramolecular structures. *J. Am. Chem. Soc.*, 124(41):12144–12153, 2002.

[4] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter. *Molecular Biology of the Cell.* Garland Science, New York, fourth edition, 2002.

[5] T. Kirchhausen. Clathrin. *Annual Review of Biochemistry*, 69:699–727, 2000.

[6] C. Smith. Two geometric solutions to a transporting problem. *Science*, 311:182–183, 2006.

[7] H. Lodish, D. Baltimore, A. Berk, S. L. Zipursky, P. Matsudaira, and J. Darnell. *Molecular Cell Biology.* Scientific American Books, New York, NY, 1995.

[8] D. E. Wakeham, J. A. Ybe, F. M. Brodsky, and P. K. Hwang. Molecular structures of proteins involved in vesicle coat formation. *Traffic*, 1:393–398, 2000.

[9] E. Kocsis, B. L. Trus, C. J. Steer, M. E. Bisher, and A. C. Steven. Image averaging of flexible fibrous macromolecules - the clathrin triskelion has an elastic proximal segment. *Journal of Structural Biology*, 107(1):6–14, 1991.

[10] A. J. Jin and R. Nossal. Rigidity of triskelion arms and clathrin nets. *Biophysical Journal*, 78(3):1183–1194, 2000.

[11] E. ter Haar, A. Musacchio, S. C. Harrison, and T. Kirchhausen. Atomic structure of clathrin: a beta propeller terminal domain joins an alpha zigzag linker. *Cell*, 95(4):563–73, 1998.

[12] J. Heuser. Cover. In *Conference Schedule*. American Society for Cell Biology, Year Unknown.

[13] J. Heuser and T. Kirchhausen. Deep-etch views of clathrin assemblies. *J Ultrastruct Res*, 92(1-2):1–27., 1985.

[14] R. A. Crowther and B. M. F. Pearse. Assembly and packing of clathrin into coats. *Journal of Cell Biology*, 91(3):790–797, 1981.

[15] W. Barouch, K. Prasad, L. E. Greene, and E. Eisenberg. Atpase activity associated with the uncoating of clathrin baskets by hsp70. *J. Biol. Chem.*, 269(46):28563–8., 1994.

[16] A. Fotin, Y. F. Cheng, P. Sliz, N. Grigorieff, S. C. Harrison, T. Kirchhausen, and T. Walz. Molecular model for a complete clathrin lattice from electron cryomicroscopy. *Nature*, 432(7017):573–579, 2004.

[17] T. Kirchhausen, S. C. Harrison, and J. Heuser. Configuration of clathrin trimers - evidence from electron-microscopy. *Journal of Ultrastructure and Molecular Structure Research*, 94(3):199–208, 1986.

[18] R. J. Nossal and J. Zimmerberg. Endocytosis: Curvature to the enth degree. *Current Biology*, 12:R770–R772, 2002.

[19] M. G. J. Ford, I. G. Mils, B. J. Peter, Y. Vallis, G. J. K. Praefcke, P. R. Evans, and H. T. McMahon. Curvature of clathrin-coated pits driven by epsin. *Nature*, 419:361–366, 2002.

[20] L. Greene, 2005. Personal Communication.

[21] E. Lafer and K. Prasad, 2004. Personal Communication.

[22] J. A. Ybe, F. M. Brodsky, K. Hofmann, K. Lin, S. H. Liu, L. Chen, T. N. Earnest, R. J. Fletterick, and P. K. Hwang. Clathrin self-assembly is mediated by a tandemly repeated superhelix. *Nature*, 399(6734):371–5, 1999.

[23] F. M. Brodsky, C. Y. Chen, C. Knuehl, M. C. Towler, and D. E. Wakeham. Biological basket weaving: formation and function of clathrin-coated vesicles. *Annu. Rev. Cell Dev. Biol.*, 17:517–68, 2001.

[24] J. E. Heuser, J. H. Keen, L. M. Amende, R. E. Lippoldt, and K. Prasad. Deep-etch visualization of 27s clathrin: A tetrahedral tetramer. *J. Cell Biol.*, 105:1999–2009, 1987.

[25] J. D. Jackson. *Classical Electrodynamics*. John Wiley and Sons, Inc, New York, 1999.

[26] Charles Tanford. *Physical Chemistry of Macromolecules*. John Wiley and Sons, Inc., New York, 1961.

[27] L. Van Hove. Correlations in space and time and born approximation scattering in systems of interaction particles. *Physical Review*, 95(1):249, 1954.

[28] Roger Pynn. Neutron scattering - a primer. *Los Alamos Science*, 1990.

[29] Gordon Baym. *Lectures on Quantum Mechanics*. Perseus Books, Cambridge, MA, 1990.

[30] C. R. Cantor and P. R. Schimmel. *Biophysical Chemistry*. W. H. Freeman and Company, San Fransisco, CA, 1980.

[31] O. Glatter and O. Kratky. *Small Angle X-ray Scattering*. Academic Press, New York, 1982.

[32] D. B. Hand. The refractivity of protein solutions. *J. Biol. Chem.*, 108(3):703, 1934.

[33] P. Debye. *Ann. Physik*, 46:809–823, 1915.

[34] B. H. Zimm, R. S. Stein, and P. Doty. Classical theory of light scattering from solutions – a review. *Polymer Bulletin*, 1(4/5):90–119., 1945.

[35] D. McIntyre and F. Gornick. *Light Scattering from Dilute Polymer Solutions*. Gordan and Breach, New York, 1964.

[36] P. Debye and P. P. Debye. *J. Applied Phys.*, 15:338, 1944.

[37] A. Guinier. *Ann. Phys.*, 12:161, 1939.

[38] J. G. Kirkwood and J. Riseman. The intrinsic viscosities and diffusion constants of flexible macromolecules in solution. *J. Chem. Phys.*, 16:565–573, 1948.

[39] V. Bloomfield, W. O. Dalton, and K. E. Van Holde. Frictional coefficients of multisubunit structures i. theory. *Biopolymers*, 5:135–148, 1967.

[40] V. Bloomfield, W. O. Dalton, and K. E. Van Holde. Frictional coefficients of multisubunit structures ii. application to proteins and viruses. *Biopolymers*, 5:149–159, 1967.

[41] C. W. Oseen. *Hydrodynamik*. Academishes Verlag, Leipzig, 1927.

[42] J. M. Burgers. On the motion of small particles of elongated form, suspended in a viscous liquid. In Committee for the Study of Viscosity of the Academy of Sciences at Amsterdam, editor, *Second Report on Viscosity and Plasticity*, pages 113–184. Nordemann Publishing, New York, NY, 1938.

[43] J. Rotne and S. Prager. Variational treatment of hydrodynamic interaction on polymers. *J. Chem. Phys.*, 50:4831–4837, 1969.

[44] H. Yamakawa. Transport properties of polymer chains in dilute solutions. hydrodynamic interaction. *J. Chem. Phys.*, 53:436–443, 1970.

[45] J. Garcia de la Torre and Bloomfield V. A. Hydrodynamic properties of macromolecular complexes i. translation. *Biopolymers*, 16:1747–1763, 1977.

[46] J. Garcia de la Torre and Bloomfield V. A. Hydrodynamic properties of macromolecular complexes ii. rotation. *Biopolymers*, 16:1764–1778, 1977.

[47] J. Garcia de la Torre and Bloomfield V. A. Hydrodynamic properties of macromolecular complexes iii. bacterial viruses. *Biopolymers*, 16:1779–1793, 1977.

[48] H. Boukari, R. Nossal, D. L. Sackett, and P. Schuck. Hydrodynamics of nanoscopic tubulin rings in dilute solutions. *Phys Rev Lett*, 93(9):098106, 2004.

[49] R. Zwanzig, J. Kiefer, and G. H. Weiss. On the validity of kirkwood-riseman theory. *Proc. Natl. Acad. Sci.*, 60:381, 1968.

[50] B. Carrasco, J. Garcia de la Torre, and P. Zipper. Calculation of hydrodynamic properties of macromolecular bead models with overlapping spheres. *Eur. Biophys. J.*, 28:510–515, 1999.

[51] J. Garcia de la Torre, S. Navarro, M. C. Lopez Martinez, F. G. Diaz, and J. J. Lopez Cascales. Hydro: a computer program for the prediction of hydrodynamic properties of macromolecules. *Biophys. J.*, 67(2):530–1., 1994.

[52] J. G. Garcia de la Torre, 2005. Personal Communication.

[53] J. G. Garcia de la Torre and V. A. Bloomfield. Hydrodynamic properties of complex, rigid, biological macromolecules: theory and applications. *Quart. Rev. Biophys.*, 14(1):81–139., 1981.

[54] J. B. Hubbard and J. F. Douglas. Hydrodynamic friction of arbitrarily shaped brownian particles. *Physical Review E*, 47(5):R2983–R2986, 1993.

[55] E. H. Kang, M. L. Mansfield, and J. F. Douglas. Numerical path integration technique for the calculation of transport properties of proteins. *Physical Review E*, 69(3), 2004.

[56] W. Matsui and T. Kirchhausen. Stabilization of clathrin coats by the core of the clathrin-associated protein complex ap-2. *Biochemistry*, P29(48):10791–10798, 1990.

[57] J. R. Morgan, K. Prasad, W. Hao, G. J. Augustine, and E. M. Lafer. A conserved clathrin assembly motif essential for synaptic vesicle endocytosis. *J. Neurosci.*, 20(23):8667–76, 2000.

[58] R. Nossal. Mechanical aspects of clathrin cage formation. *Macromolecular Symposia*, 227:17–26, 2005.

[59] P. Nandi, G. Irace, P. Van Jaarsveld, R. Lippoldt, and H. Edelhoch. Instability of coated vesicles in concentrated sucrose solutions. *Proc. Nat. Acad. Sci.*, 79:5881–5885, 1982.

[60] K. Prasad and R. Lippoldt. Molecular characterization of the ap_{180} coated vesicle assembly protein. *Biochemistry*, 27:6098–6104, 1988.

[61] C. N. Pace, F. Vajdos, L. Fee, G. Grimsley, and T. Gray. How to measure and predict the molar absorption coefficient of a protein. *Protein Sci.*, 4(11):2411–23, 1995.

[62] M. L. Ferguson, K. Prasad, D. L. Sackett, H. Boukari, E. M. Lafer, and R. Nossal.

The conformation of a clathrin triskelion in solution. *Biochemistry*, 45(19):7, 2006.

[63] B. J. Berne and R. Pecora. *Dynamic Light Scattering with Applications to Chemistry, Biology and Physics*. John Wiley and Sons, Inc., New York, 1st edition, 1976.

[64] D E Koppel. Analysis of macromolecular polydispersity in intensity correlation spectroscopy - method of cumulants. *J. Chem. Phys.*, 57(11):4814, 1972.

[65] A. Musacchio, C. J. Smith, A. M. Roseman, S. C. Harrison, T. Kirchhausen, and B. M. F. Pearse. Functional organization of clathrin in coats: Combining electron cryomicroscopy and x-ray crystallography. *Molecular Cell*, 3(6):761–770, 1999.

[66] M. M. Tirado Garcia, M. A. Jimenez Rios, and J. M. Garcia Bernal. Translational diffusion and intrinsic viscosity of globular proteins: Theoretical predictions using hydrated hydrodynamic models (application to bpti). *Int. J. Biol. Macromol.*, 12(1):19–24, 1990.

[67] T. Hellweg, W. Eimer, E. Krahn, K. Schneider, and A. Muller. Hydrodynamic properties of nitrogenase – the mofe protein from *azobacter vivelandii* studied by dynamic light scattering and hydrodynamic modelling. *Biochimica et Biophysica Acta*, 1337:311–318, 1997.

[68] T. Yoshimura, K. Kameyama, S. Maezawa, and T. Takagi. Skeletal structure

of clathrin triskelion in solution: experimental and theoretical approaches. *Biochemistry*, 30(18):4528–34., 1991.

[69] S. Zaremba and J. H. Keen. Assembly polypeptides from coated vesicles mediate reassembly of unique clathrin coats. *Journal of Cell Biology*, 97(5):1339–1347, 1983.

[70] R. Nossal. Energetics of clathrin basket assembly. *Traffic*, 2(2):138–47, 2001.

[71] S. H. Liu, M. L. Wong, C. S. Craik, and F. M. Brodsky. Regulation of clathrin assembly and trimerization defined using recombinant triskelion hubs. *Cell*, 83(2):257–67, 1995.

[72] J. A. Ybe, B. Greene, S. H. Liu, U. Pley, P. Parham, and F. M. Brodsky. Clathrin self-assembly is regulated by three light-chain residues controlling the formation of critical salt bridges. *EMBO J.*, 17(5):1297–303, 1998.

[73] C. M. Marques and G. H. Fredrickson. Rigid gaussian chains i: the scattering function. *J. Phys. II France*, 7:1805–1816, 1997.

[74] D. L. Sackett, V. Chernomordik, S. Krueger, and R. Nossal. Use of small-angle neutron scattering to study tubulin polymers. *Biomacromolecules*, 4:461–467, 2003.

[75] A. I. Norman, R. Ivkov, J. G. Forbes, and S. C. Greer. The polymerization of actin: Structural changes from small-angle neutron scattering. *J. Chem. Phys.*, 123:154904, 2005.

[76] P. S. Naranjan, P. B. Yim, J. G. Forbes, S. C. Greer, J. Dudowicz, K. F. Freed, and J. F. Doublas. The polymerization of actin: Thermodynamics near the polymerization line. *J. Chem. Phys.*, 119(7):4070–4084, 2003.

[77] C. H. Cho, J. Urquidi, S. Sing, and G. W. Robinson. Thermal offset viscosities of liquid h2o, d2o, and t2o. *J. Phys. Chem. B*, 103(11):1991–1994, 1999.

[78] R. C. Hardy and R. L. Cottington. Viscosity of deuterium oxide and water from 5° to 125° c. *J. Chem. Phys.*, 17(5):509–510, 1949.

[79] M. L. Ferguson, K. Prasad, H. Boukari, D. L Sackett, S. Krueger, E. M. Lafer, and R. Nossal. Small angle neutron scattering studies of clathrin triskelia in solution show evidence of molecular flexibility. *In preparation*, 2007.

[80] D. Sackett and M. L. Ferguson. Unpublished data. 2007.

[81] I. Teraoka. *Polymer Solutions: An Introduction to Physical Properties.* Wiley-Interscience, Brooklyn, NY, 2002.

[82] R. A. Harris and J. E. Hearst. On polymer dynamics. *J. Chem. Phys.*, 44(7):2595–2602, 1966.

[83] F. Reif. *Fundamentals of statistical and thermal physics.* McGraw-Hill Book Company, New York, 1964.

[84] H. Fischer, I. Polikarpov, and A. F. Craievich. Average protein density is a molecular-weight-dependent function. *Protein Sci.*, 13:2825–2828, 2004.

[85] J. Howard. *Mechanics of Motor Proteins and the Cytoskeleton.* Sinauer Associates, Inc., Sunderland, MA, 2001.

[86] F. Gittes, B. Mickey, J. Nettleton, and J. Howard. Flexural rigidity of microtubules and actin filaments measured from thermal fluctuations in shape. *J. Cell Biol.*, 120(4):923–934, 1993.

[87] R. W. Pastor. Techniques and applications of langevin dynamic simulations. In G. R. Luckhurst and C. A. Veracini, editors, *The Molecular Dynamics of Liquid Crystals*, pages 85–138. Klewer Academic Publishers, Amsterdam, NE, 1994.

[88] R. Skupsky, 2006. Personal Communication.

[89] R. W. Pastor and M. Karplus. Parametrization of the friction constant for stochastic simulations of polymers. *J. Phys. Chem.*, 92(9):2636–2641, 1988.

[90] D. Harries, 2006. Personal Communication.

[91] D. Ermak and J. A. McCammon. Brownian dynamics with hydrodynamic interactions. *J. Chem. Phys.*, 69(4):1352–1360, 1978.

[92] R. Pastor, 2006. Personal Communication.

[93] K. Y. Sanbonmatsu, S. Joseph, and C. S. Tung. Simulating movement of trna into the ribosome during decoding. *Proc. Nat. Acad. Sci.*, 102(44):15854–9, 2005.

[94] Charles L. Asbury, Adrian N. Fehr, and Steven M. Block. Kinesin Moves by an Asymmetric Hand-Over-Hand Mechanism. *Science*, 302(5653):2130–2134, 2003.

[95] R. Yasuda, H. Noji, K. Kinosita, and M. Yoshida. F1-atpase is a highly efficient molecular motor that rotates with discrete 120 degree steps. *Cell*, 93(7):1117–1124, 1998.

[96] Albert J. Jin, Kondury Prasad, Paul D. Smith, Eileen M. Lafer, and Ralph Nossal. Measuring the elasticity of clathrin-coated vesicles via atomic force microscopy. *Biophys. J.*, 90(9):3333–3344, 2006.

[97] A. Fotin, Y. F. Cheng, N. Grigorieff, T. Walz, S. C. Harrison, and T. Kirchhausen. Molecular model for a complete clathrin lattice from electron cryomicroscopy. *Nature*, 432:649–653, 2004.

[98] J.B. Heymann, K. Iwasaki, Y. I. Yim, N. Cheng, D.M. Belnap, L.E. Greene, E. Eisenberg, and A. C. Steven. Visualization of the binding of hsc70 atpase to clathrin baskets: implications for an uncoating mechanism. *J. Biol. Chem.*, 280(8):7156–7161, 2005.